Designation: ~~E2919 – 14~~ E2919 – 22

Standard Test Method for Evaluating the Performance of Systems that Measure Static, Six Degrees of Freedom (6DOF), Pose[1]

This standard is issued under the fixed designation E2919; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 *Purpose*—In this test method, metrics and procedures for collecting and analyzing data to determine the performance of a pose measurement system in computing the pose (position and orientation) of a rigid object are provided.

1.2 This test method applies to the situation in which both the object and the pose measurement system are static with respect to each other when measurements are performed. Vendors may use this test method to establish the performance limits for their six degrees of freedom (6DOF) pose measurement systems. The vendor may use the procedures described in 9.2 to generate the test statistics, then apply an appropriate margin or scaling factor as desired to generate the performance specifications. This test method also provides a uniform way to report the relative or absolute pose measurement capability of the system, or both, making it possible to compare the performance of different systems.

1.3 *Test Location*—The methodology defined in this test method shall be performed in a facility in which the environmental conditions are within the pose measurement system's rated conditions and meet the user's requirements.

1.4 *Units*—The values stated in SI units are to be regarded as the standard. No other units of measurement are included in this standard.

1.5 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.6 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E456 Terminology Relating to Quality and Statistics

E2544 Terminology for Three-Dimensional (3D) Imaging Systems

2.2 *ANSI/NCSL Standard:*[3]

ANSI/NCSL Z540.3:2006 Requirements for the Calibration of Measuring and Test Equipment

---

[1] This test method is under the jurisdiction of ASTM Committee E57 on 3D Imaging Systems and is the direct responsibility of Subcommittee E57.23 on Industrial 3D Machine Vision Systems.

Current edition approved ~~July 1, 2014~~Dec. 15, 2022. Published ~~August 2014~~February 2023. Originally approved in 2013. Last previous edition approved in ~~2013~~2014 as ~~E2919~~E2919 – 14.~~-13.~~ DOI: ~~10.1520/E2919-14.~~10.1520/E2919-22.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

2.3 *ASME Standard:*[4]

ASME B89.4.19 Performance Evaluation of Laser-Based Spherical Coordinate Measurement Systems

2.4 *ISO/IEC Standards:*[5]

JCGM 100:2008 Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM)
JCGM 106:2012 Evaluation of measurement data – The role of measurement uncertainty in conformity assessment
JCGM 200:2012 International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM), 3rd edition
~~JCGM 100:2008 Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM)~~
IEC 60050-300:2001 International Electrotechnical Vocabulary—Electrical and Electronic Measurements and Measuring Instruments

3. Terminology

3.1 *Definitions from Other Standards:*

3.1.1 *calibration, n*—operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication. **JCGM 200:2012**

3.1.1.1 *Discussion*—

(1) A calibration may be expressed by a statement, calibration function, calibration diagram, calibration curve, or calibration table. In some cases, it may consist of an additive or multiplicative correction of the indication with associated measurement uncertainty.

(2) Calibration should not be confused with either adjustment of a measuring system, often mistakenly called "self-calibration," or verification of calibration.

(3) Often, the first step alone in 3.1.1 is perceived as being calibration.

3.1.2 *maximum permissible measurement error, maximum permissible error, and limit of error, n*—extreme value of measurement error, with respect to a known reference quantity value, permitted by specifications or regulations for a given measurement, measuring instrument, or measuring system. **JCGM 200:2012**

3.1.2.1 *Discussion*—

(1) Usually, the terms "maximum permissible errors" or "limits of error" are used when there are two extreme values.

(2) The term "tolerance" should not be used to designate "maximum permissible error."

3.1.3 *measurand, n*—quantity intended to be measured. **JCGM 200:2012**

3.1.3.1 *Discussion*—

(1) The specification of a measurand requires knowledge of the kind of quantity; description of the state of the phenomenon, body, or substance carrying the quantity, including any relevant component; and the chemical entities involved.

(2) In the second edition of the VIM and IEC 60050-300, the measurand is defined as the "quantity subject to measurement."

(3) The measurement, including the measuring system and the conditions under which the measurement is carried out, might change the phenomenon, body, or substance such that the quantity being measured may differ from the measurand as defined. In this case, adequate correction is necessary.

(a) Example 1—The potential difference between the terminals of a battery may decrease when using a voltmeter with a significant internal conductance to perform the measurement. The open-circuit potential difference can be calculated from the internal resistances of the battery and the voltmeter.

(b) Example 2—The length of a steel rod in equilibrium with the ambient Celsius temperature of 23°C will be different from the length at the specified temperature of 20°C, which is the measurand. In this case, a correction is necessary.

(4) In chemistry, "analyte," or the name of a substance or compound, are terms sometimes used for "measurand." This usage is erroneous because these terms do not refer to quantities.

3.1.4 *measurement error, error of measurement, and error, n*—measured quantity value minus a reference quantity value. **JCGM 200:2012**

3.1.4.1 *Discussion*—

(1) The concept of "measurement error" can be used both:

---

[4] Available from American Society of Mechanical Engineers (ASME), ASME International Headquarters, Three Park Ave., New York, NY 10016-5990, http://www.asme.org.

[5] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

(a) When there is a single reference quantity value to refer to, which occurs if a calibration is made by means of a measurement standard with a measured quantity value having a negligible measurement uncertainty or if a conventional quantity value is given, in which case the measurement error is known, and

(b) If a measurand is supposed to be represented by a unique true quantity value or a set of true quantity values of negligible range, in which case the measurement error is not known.

(2) Measurement error should not be confused with production error or mistake.

3.1.5 *measurement sample and sample, n*—group of observations or test results, taken from a larger collection of observations or test results, that serves to provide information that may be used as a basis for making a decision concerning the larger collection. **E456**

3.1.6 *measurement uncertainty, uncertainty of measurement, and uncertainty, n*—non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand based on the information used. **JCGM 200:2012**

3.1.6.1 *Discussion—*

(1) Measurement uncertainty includes components arising from systematic effects, such as components associated with corrections and the assigned quantity values of measurement standards, as well as the definitional uncertainty. Sometimes estimated systematic effects are not corrected for but, instead, associated measurement uncertainty components are incorporated.

(2) The parameter may be, for example, a standard deviation called standard measurement uncertainty (or a specified multiple of it) or the half width of an interval, having a stated coverage probability.

(3) Measurement uncertainty comprises, in general, many components. Some of these may be evaluated by Type A evaluation of measurement uncertainty from the statistical distribution of the quantity values from series of measurements and can be characterized by standard deviations. The other components, which may be evaluated by Type B evaluation of measurement uncertainty, can also be characterized by standard deviations evaluated from probability density functions based on experience or other information.

(4) In general, for a given set of information, it is understood that the measurement uncertainty is associated with a stated quantity value attributed to the measurand. A modification of this value results in a modification of the associated uncertainty.

3.1.7 *precision, n*—closeness of agreement between independent test results obtained under stipulated conditions. **E456**

3.1.7.1 *Discussion—*

(1) Precision depends on random errors and does not relate to the true value or the specified value.

(2) 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.

(3) "Independent test results" means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular sets of extreme stipulated conditions.

3.1.8 *rated conditions, n*—manufacturer-specified limits on environmental, utility, and other conditions within which the manufacturer's performance specifications are guaranteed at the time of installation of the instrument. **ASME B89.4.19**

3.1.9 *reference quantity value and reference value, n*—quantity value used as a basis for comparison with values of quantities of the same kind. **JCGM 200:2012**

3.1.9.1 *Discussion—*

(1) A reference quantity value can be a true quantity value of a measurand, in which case it is unknown, or a conventional quantity value, in which case it is known.

(2) A reference quantity value with associated measurement uncertainty is usually provided with reference to:

(a) A material, for example, a certified reference material;

(b) A device, for example, a stabilized laser;

(c) A reference measurement procedure; and

(d) A comparison of measurement standards.

3.1.10 *registration, n*—process of determining and applying to two or more datasets the transformations that locate each dataset in a common coordinate system so that the datasets are aligned relative to each other. **E2544**

3.1.10.1 *Discussion—*

(1) A three-dimensional (3D) imaging system generally collects measurements in its local coordinate system. When the same scene or object is measured from more than one position, it is necessary to transform the data so that the datasets from each position have a common coordinate system.

(2) Sometimes the registration process is performed on two or more datasets that do not have regions in common. For example, when several buildings are measured independently, each dataset may be registered to a global coordinate system instead of to each other.

(3) In the context of this definition, a dataset may be a mathematical representation of surfaces or may consist of a set of coordinates, for example, a point cloud, a 3D image, control points, survey points, or reference points from a computer-aided drafted (CAD) model. Additionally, one of the datasets in a registration may be a global coordinate system (as in 3.1.10.1*(2)*).

(4) The process of determining the transformation often involves the minimization of an error function, such as the sum of the squared distances between features (for example, points, lines, curves, and surfaces) in two datasets.

(5) In most cases, the transformations determined from a registration process are rigid body transformations. This means that the distances between points within a dataset do not change after applying the transformations, that is, rotations and translations.

(6) In some cases, the transformations determined from a registration process are nonrigid body transformations. This means that the transformation includes a deformation of the dataset. One purpose of this type of registration is to attempt to compensate for movement of the measured object or deformation of its shape during the measurement.

(7) Registration between two point clouds is sometimes referred to as cloud-to-cloud registration, between two sets of control or survey points as target-to-target, between a point cloud and a surface as cloud-to-surface, and between two surfaces as surface-to-surface.

(8) The word alignment is sometimes used as a synonymous term for registration. However, in the context of this definition, an alignment is the result of the registration process.

3.1.11 *true quantity value, true value of a quantity, and true value, n*—quantity value consistent with the definition of a quantity.

JCGM 200:2012

3.1.11.1 *Discussion*—

(1) In the error approach to describing measurement, a true quantity value is considered unique and, in practice, unknowable. The uncertainty approach is to recognize that, owing to the inherently incomplete amount of detail in the definition of a quantity, there is not a single true quantity value but rather a set of true quantity values consistent with the definition. However, this set of values is, in principle and practice, unknowable. Other approaches dispense altogether with the concept of true quantity value and rely on the concept of metrological compatibility of measurement results for assessing their validity.

(2) In the special case of a fundamental constant, the quantity is considered to have a single true quantity value.

(3) When the definitional uncertainty associated with the measurand is considered to be negligible compared to the other components of the measurement uncertainty, the measurand may be considered to have an "essentially unique" true quantity value. This is the approach taken by JCGM 100 and associated documents in which the word "true" is considered to be redundant.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *absolute pose, n*—pose of an object in the coordinate frame of the system under test.

3.2.2 *degree of freedom, DOF, n*—any of the minimum number of translation or rotation components required to specify completely the pose of a rigid body.

3.2.2.1 *Discussion*—

(1) In a 3D space, a rigid object can have at most 6DOFs, three translation and three rotation.

(2) The term "degree of freedom" is also used with regard to statistical testing. It will be clear from the context in which it is used whether the term relates to a statistical test or the rotation/translation aspect of the object.

3.2.3 *pose, n*—a 6DOF vector whose components represent the position and orientation of a rigid object with respect to a coordinate frame.

3.2.4 *pose measurement system, n*—a 3-D imaging system that measures the pose of an object.

3.2.4.1 *Discussion*—

This system can consist of both hardware and software.

3.2.5 *reference system, n*—a measurement instrument or system used to generate a reference value or quantity.

3.2.6 *relative pose, n*—change of an object's pose between two poses measured in the same coordinate frame.

3.2.7 *system under test, SUT, n*—measurement instrument or system used to generate a test value or quantity.

3.2.8 *work volume, n*—physical space, or region within a physical space, that defines the bounds within which a pose measurement system is acquiring data.

3.3 *Notation:*

3.3.1 Mathematical equations throughout this test method use the following notation. Scalar variables are lower-cased italicized (for example, x), and scalar constants are upper-case and italicized (for example, N). Vectors are lower-case and bold faced (for example, $\mathbf{t}$), and matrices are upper-case and bold-faced (for example, $\mathbf{H}$). Special characters are used to denote the measurements from the system under test (SUT). The hat symbol (for example, $\hat{\mathbf{R}}$) represents a measurement from the SUT in its own coordinate frame, while the tilde (for example, $\tilde{\mathbf{R}}$) represents a measurement from the reference system (RS) coordinate frame transformed to the SUT system coordinate frame.

4. Summary of Test Method

4.1 This test method provides a set of test procedures and statistically based performance metrics to evaluate quantitatively the performance of a 6DOF pose measurement system to measure the static pose of an object. It is applicable to the situation in which both the pose measurement system and the object are static with respect to each other when the measurements are performed. The test method allows for the evaluation of the absolute and relative pose of an object.

4.2 The test method involves measuring the pose of a user-specified object with the SUT and an RS at a minimum of 32 random locations within the work volume of the SUT. The pose errors, absolute or relative, are calculated based on the measurements from the SUT and the RS. Performance of the SUT with regard to the vendor's specifications pertaining to the user's application is determined by selecting the appropriate statistical test or tests as determined by the user.

5. Significance and Use

5.1 Pose measurement systems are used in a wide range of fields including manufacturing, material handling, construction, medicine, and aerospace. The use of pose measurement systems could, for example, replace the need to fix the poses of objects of interest by mechanical means.

5.2 Potential users have difficulty comparing pose measurement systems because of the lack of standard performance specifications and test methods, and must rely on the specifications of a vendor regarding the system's performance, capabilities, and suitability for a particular application. This standard makes it possible for a user to assess and compare the performance of candidate pose measurement systems, and allows the user to determine if the measured performance results are within the vendor's claimed specifications with regard to the user's application. This standard also facilitates the improvement of pose measurement systems by providing a common set of metrics to evaluate system performance.

5.3 The intent of this test method is to allow a user to determine the performance of a vendor's system under conditions specific to the user's application, and to determine whether the system still performs in accordance with the vendor's specifications under those conditions. The intention of this test method is not to validate a vendor's claims; although, under specific situations, this test method may be adapted for this purpose.

6. Apparatus

6.1 *Reference System:*

6.1.1 A reference pose measurement shall be established so that the error of the measured pose can be evaluated. If possible, the pose measurement uncertainty associated with the RS should be an order of magnitude (ten times) less than the measurement uncertainty associated with the SUT based on the vendor's specifications. The RS shall have been calibrated within the vendor-recommended calibration cycle and reported as described in Section 11. The RS shall have been calibrated according to an available published standard. For example, laser trackers or coordinate measurement machines that comply with ASME B89 can be used to obtain the reference values.

6.2 *Test Objects:*

6.2.1 Test objects should be rigid bodies chosen based on the user's intended purpose or application. The geometry of the objects should be representative of the user's application; if the user has no specific application, simple object geometries designed to minimize or eliminate pose ambiguities can be used. See English (**1**)[6] for an illustrative example of a possible geometric test object designed to minimize pose ambiguity.

6.2.2 In this test method, no restrictions on the properties of the selected test objects (for example, material, size, reflectivity, or texture) are placed; however, user or vendor restrictions on the test object's properties may need to be accommodated if using this test method to evaluate the performance of the system with regard to the vendor's specifications as they pertain to the user's application.

7. Sampling Size

7.1 The performance evaluation of the SUT is based on the measurement error of a set of measurement results. The set consists of randomly sampled data points obtained from within the work volume. Assuming that any single measurement error depends only on the pose being measured, and not on the sequence of poses measured, the sample size $N \geq 32$, should ensure that the average error approaches a normal distribution per the Central Limit Theorem.

8. Absolute Pose Error and Relative Pose Error

8.1 This section describes methods for calculating the absolute and relative pose errors. The concepts of absolute and relative pose error will be explained in greater detail in 8.2 and 8.3, respectively. These errors form the basis for the test procedure discussed in Section 9, which will then be used for the performance evaluations in Section 10.

8.1.1 Consider an instrument, *S*, performing pose measurements of an object, *O*, at Pose $k = 1, 2, \ldots, N$. The pose consists of both orientation and position information. This test method uses a 3 × 3 rotation matrix to represent rotation and a 3 × 1 vector to represent translation. Methods for transforming other rotation representations into a 3 × 3 rotation matrix representation can be found in Huynh (**2**).

8.1.2 The rotation and translation information at Pose *k* can be simultaneously represented as a 4 × 4 homogeneous matrix.

$$ {}_{S}\mathbf{H}_{O_k}=\begin{bmatrix}\mathbf{R}_k & \mathbf{t}_k \\ 0 & 1\end{bmatrix} \tag{1}$$

8.1.2.1 Here, the 3 × 3 rotation matrix, $\mathbf{R}_k$, represents the rotation of the object, *O*, in the coordinate system of *S* at Pose *k* and $\mathbf{t}_k$ represents the 3 × 1 translation vector of the object, *O*, in the coordinate system of *S* at Pose *k*.

8.1.3 In 8.2 and 8.3, methods are described to evaluate the SUT with respect to a RS. In this test method, the poses of the SUT and the RS are fixed relative to each other; therefore, there is a rigid transformation between them. Here,

$$ {}_{SUT}\hat{\mathbf{H}}_{O_k}=\begin{bmatrix}\hat{\mathbf{R}}_k & \hat{\mathbf{t}}_k \\ 0 & 1\end{bmatrix} \tag{2}$$

represents the object pose in the coordinate frame of the SUT at Pose *k* and

$$ {}_{RS}\mathbf{H}_{O_k}=\begin{bmatrix}\mathbf{R}_k & \mathbf{t}_k \\ 0 & 1\end{bmatrix} \tag{3}$$

represents the object pose in the coordinate frame of the RS at Pose *k*. In 8.2, a method is described to calculate the absolute pose error of the object in a common coordinate frame. In 8.3, a method is described to calculate the relative pose error of the object in which the SUT relative pose is calculated in the SUT coordinate frame and the RS relative pose is calculated in the RS coordinate frame.

8.2 *Absolute Pose Error:*

8.2.1 The absolute pose is defined with respect to the coordinate frame of the SUT. As a result, the object pose in the coordinate frame of the RS shall be transformed to the coordinate frame of the SUT. It is assumed that the coordinate frames of the RS and the SUT are fixed relative to one another and, therefore, the transformation between their respective coordinate frames does not change. The RS shall be registered to the SUT according to the vendor's specified process. In the case that the vendor does not provide means for registration, the selection of methods for transforming the coordinate frame is left to the user. Note that the

---

[6] The boldface numbers in parentheses refer to the list of references at the end of this standard.

registration process contributes toward the total measurement error (see 9.1.2). Once transformed, the absolute pose of the object computed from the measurement results obtained from the RS can be compared with the absolute pose of the object computed from the measurement results obtained from the SUT to determine the rotation measurement error and the translation measurement error.

8.2.2 Here, the absolute pose of an object at Pose *k* computed from the measurement results obtained from the RS is represented as:

$$
\begin{aligned}
{}_{SUT}\tilde{\mathbf{H}}_{O_k} &= {}_{SUT}\mathbf{H}_{RS} \times {}_{RS}\mathbf{H}_{O_k} \\
&= \begin{bmatrix} {}_{SUT}\mathbf{R}_{RS} & {}_{SUT}\mathbf{t}_{RS} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_k & \mathbf{t}_k \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \tilde{\mathbf{R}}_k & \tilde{\mathbf{t}}_k \\ 0 & 1 \end{bmatrix}
\end{aligned} \tag{4}
$$

where:

${}_{SUT}\mathbf{H}_{RS}$ = transformation matrix of the coordinate frame of the RS to the SUT (see Fig. 1), and

$$
\tilde{\mathbf{R}}_k = {}_{SUT}\mathbf{R}_{RS}\mathbf{R}_k \tag{5}
$$

$$
\begin{aligned}
\tilde{\mathbf{t}}_k &= {}_{SUT}\mathbf{R}_{RS}\mathbf{t}_k + {}_{SUT}\mathbf{t}_{RS} \\
&= [\tilde{x}_k \quad \tilde{y}_k \quad \tilde{z}_k]^{\mathrm{T}}
\end{aligned}
$$

are the rotation and translation components of the absolute pose computed from the measurement results obtained from the RS at Pose *k* in the SUT coordinate frame.

8.2.3 The absolute pose of an object at Pose *k* computed from the SUT is represented as:

$$
{}_{SUT}\hat{\mathbf{H}}_{O_k} = \begin{bmatrix} \hat{\mathbf{R}}_k & \hat{\mathbf{t}}_k \\ 0 & 1 \end{bmatrix} \tag{6}
$$

where:

$\hat{\mathbf{R}}_k$ = rotation component of the absolute pose computed from the SUT at Pose *k*, and

$\hat{\mathbf{t}}_k$ = $[\hat{x}_k \; \hat{y}_k \; \hat{z}_k]^{\mathrm{T}}$ = translation component of the absolute pose computed from the SUT at Pose *k*.

8.2.4 Using this notation, the rotation measurement error can be computed using the following procedure:

8.2.4.1 Compute ${}_{SUT}\mathbf{R}_{RS}$ from ${}_{SUT}\mathbf{H}_{RS}$.

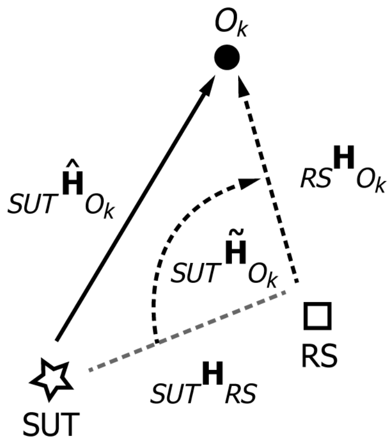

FIG. 1 Absolute Pose of Object *O* at Pose *k* Computed from the SUT Represented by ${}_{SUT}\hat{\mathbf{H}}_{O_k}$ and Computed from the RS Represented by ${}_{SUT}\tilde{\mathbf{H}}_{O_k} = {}_{SUT}\mathbf{H}_{RS} \times {}_{RS}\mathbf{H}_{O_k}$

8.2.4.2 8.2.4.2 8.2.4.2 8.2.4.2 Transform the orientation data obtained from the RS into the coordinate frame of the SUT by $\tilde{\mathbf{R}}_k = {}_{SUT}\mathbf{R}_{RS}\mathbf{R}_k$.

8.2.4.3 Compute the rotation difference, $\mathbf{R}_k = \tilde{\mathbf{R}}_k \hat{\mathbf{R}}_k^T$. Note that if $\tilde{\mathbf{R}}_k$ and $\hat{\mathbf{R}}_k$ are identical, then $\mathbf{R}_k$ will equal the identity matrix.

8.2.4.4 Compute the rotation measurement error as:

$$0 \le e_{\text{AbsAngle},k} = \cos^{-1}\left(\frac{\text{trace}(\mathbf{R}_k) - 1}{2}\right) < \pi \tag{7}$$

or

$e_{\text{AbsRoll,k}}$ = roll($\mathbf{R}_k$) = rotation angle error about the x axis
$e_{\text{AbsPitch,k}}$ = pitch($\mathbf{R}_k$) = rotation angle error about the y axis
$e_{\text{AbsYaw,k}}$ = yaw($\mathbf{R}_k$) = rotation angle error about the z axis

as defined in Jazar (**3**).

8.2.5 The translation measurement errors can be evaluated as follows:

$$e_{\text{AbsTran},k} = \sqrt{(\hat{x}_k - \tilde{x}_k)^2 + (\hat{y}_k - \tilde{y}_k)^2 + (\hat{z}_k - \tilde{z}_k)^2} \tag{8}$$

$$e_{\text{AbsX},k} = \hat{x}_k - \tilde{x}_k$$

$$e_{\text{AbsY},k} = \hat{y}_k - \tilde{y}_k$$

$$e_{\text{AbsZ},k} = \hat{z}_k - \tilde{z}_k$$

8.3 *Relative Pose Error:*

8.3.1 The relative pose is defined as the change of an object's pose between two poses, *j* and *k*, in the same coordinate frame. In this test method, Pose *j* is the first sample pose, while Pose *k* is selected from the remaining set of sample Poses 2 to *N*. The relative pose as seen by the SUT is compared with the relative pose as seen by the RS (see Fig. 2). The relative pose metric consists of two error components: the rotation measurement error and the translation measurement error.

8.3.2 The relative pose between Pose 1 and Pose *k* as seen by the SUT can be defined as:

$$_{O_1}\hat{\mathbf{H}}_{O_k} = {}_{SUT}\hat{\mathbf{H}}_{O_1}^{-1} \times {}_{SUT}\hat{\mathbf{H}}_{O_k} \tag{9}$$

$$= \begin{bmatrix} \hat{\mathbf{R}}_1 & \hat{\mathbf{t}}_1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{R}}_k & \hat{\mathbf{t}}_k \\ 0 & 1 \end{bmatrix}$$

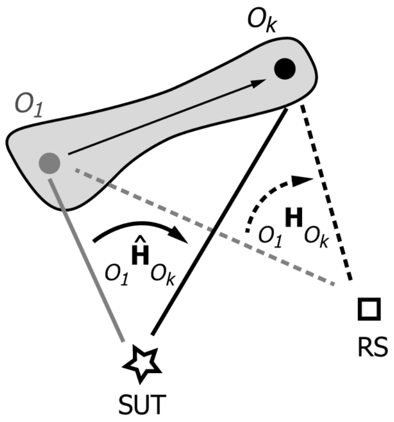

FIG. 2 Relative Pose in which Object *O* is Moving from Pose 1 to Pose *k* with Respect to the RS, which is Represented by $_{O_1}H_{O_k}$, and the SUT, which is Represented by $_{O_1}\hat{H}_{O_k}$, and the Gray Region Represents the Volume in which the Object is Being Moved from Pose O_1 to O_k

$$=\begin{bmatrix} {}_1\hat{\mathbf{R}}_k & {}_1\hat{\mathbf{t}}_k \\ 0 & 1 \end{bmatrix}$$

while the relative pose between Pose 1 and Pose k as seen by the RS can be defined as:

$$ {}_{O_1}\mathbf{H}_{O_k} = {}_{RS}\mathbf{H}_{O_1}^{-1} \times {}_{RS}\mathbf{H}_{O_k} \tag{10}$$

$$=\begin{bmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{R}_k & \mathbf{t}_k \\ 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} {}_1\mathbf{R}_k & {}_1\mathbf{t}_k \\ 0 & 1 \end{bmatrix}$$

8.3.3 The rotation measurement error can be evaluated in the following way:

8.3.3.1 Compute the rotation change as seen by the SUT from Pose 1 to Pose k as the rotation matrix, ${}_1\hat{\mathbf{R}}_k=\hat{\mathbf{R}}_1^T\hat{\mathbf{R}}_k$, and from Pose 1 to Pose k as seen by the RS as ${}_1\mathbf{R}_k=\mathbf{R}_1^T\mathbf{R}_k$.

8.3.3.2 Compute the rotation difference matrix, $\mathbf{R}_k={}_1\mathbf{R}_k\hat{\mathbf{R}}_k^T$.

8.3.3.3 Compute the rotation measurement error as:

$$0 \le e_{\mathrm{RelAngle},k} = \cos^{-1}\left(\frac{\mathrm{trace}(\mathbf{R}_k) - 1}{2}\right) < \pi \tag{11}$$

or

$e_{\mathrm{RelRoll,k}}$ = roll($\mathbf{R}_k$) = rotation angle error about the x axis
$e_{\mathrm{RelPitch,k}}$ = pitch($\mathbf{R}_k$) = rotation angle error about the y axis
$e_{\mathrm{RelYaw,k}}$ = yaw($\mathbf{R}_k$) = rotation angle error about the z axis

as defined in Jazar (3).

8.3.4 Translation measurement error can be evaluated by calculating:

$$e_{\mathrm{RelTran},k} = \sqrt{(\hat{x}_k - \hat{x}_1)^2 + (\hat{y}_k - \hat{y}_1)^2 + (\hat{z}_k - \hat{z}_1)^2} - \sqrt{(x_k - x_1)^2 + (y_k - y_1)^2 + (z_k - z_1)^2} \tag{12}$$

where:

$$\hat{\mathbf{t}}_k = [\hat{x}_k \;\; \hat{y}_k \;\; \hat{z}_k \;]^{\mathrm{T}} \tag{13}$$

= translation component of the object at Pose k as seen by the SUT, and

$$\mathbf{t}_k = [x_k \;\; y_k \;\; z_k \;]^{\mathrm{T}} \tag{14}$$

= translation component of the object at Pose k as seen by the RS.

9. Procedure

9.1 *Introduction:*

9.1.1 In this section, the basic procedure is described to determine the pose measurement error of a pose measurement system. This procedure provides the basis for the evaluation of a pose measurement system that measures the 6DOF pose of an object by comparing the results from a SUT to the results obtained from a RS.

9.1.2 The pose measurement performance can be affected by many non-system parameters and factors, including those listed in Section 11. The performance of a pose measurement system can also be affected by other factors such as those listed in 9.1.2.1 through 9.1.2.3. These errors should be minimized as much as possible.

9.1.2.1 *Noise*—Active equipment in the same environment as the pose measurement system may create noise that interferes (for example, electromagnetic noise) with the performance of the pose measurement system. Environmental factors may introduce noise that may also affect the performance of the pose measurement system.

9.1.2.2 *Registration Error*—Registration processes contribute toward the final measurement error, and the magnitude of the registration error may differ depending on the registration method used.

9.1.2.3 *Vibration*—Sensor and object vibration during the test introduces distortion into the measurement results.

9.1.3 For a given sample pose, both the RS and SUT should measure the reference object's pose simultaneously from their respective fixed poses. When testing in conditions where it is not possible for the RS and SUT to measure simultaneously, the reference measurement and the measurement from the SUT should be obtained as close together in time as possible. The SUT, RS, and reference object should not be moved during the intermittent time span until both measurements have been collected. The environmental conditions should be as consistent as possible and should be within the rated conditions of the RS and SUT over the entire period of the test.

9.2 *Test Sequence*—The basic test procedure consists of obtaining measurement results from within the work volume of the pose measurement system according to the six steps in 9.2.1 through 9.2.6. Testers may choose to either measure randomly the pose of a selected object within a user-specified subset of the work volume of the SUT (for example, a user's application may only require that poses be measured in one or more subregions of the work volume) or measure randomly the pose of the object throughout the entire work volume. The number of random pose measurements shall be as large as practical for the given SUT and RS considering the cost and complexity of acquiring pose measurements. The number of random pose measurements shall not be less than $N = 32$.

9.2.1 *Step 1*—Set up the RS and the pose measurement SUT at fixed locations according to the vendors' specifications.

9.2.2 *Step 2*—Randomly select a pose for the object. This pose will be measured by the SUT and the RS, and measurement results will consist of measured values for position (x,y, and z) and orientation (a 3 × 3 rotation matrix, **R**, see Section 8).

9.2.3 *Step 3*—Calculate the measurement errors for the translation and rotation, either absolute (Eq 7 and Eq 8, respectively) or relative (Eq 11 and Eq 12, respectively), as per Section 8 for the selected pose of the object as observed by the SUT.

9.2.4 *Step 4*—Perform Steps 2 and 3 for N sample locations within the work volume to generate a collection of measurement errors, e_1, e_2, …, e_N.

9.2.5 *Step 5*—Calculate the average measurement error, $\bar{e}$, as:

$$\bar{e} = \frac{\sum_{k=1}^{N} e_k}{N} \tag{15}$$

Compute the variance, s^2, using:

$$s^2 = \frac{\sum_{k=1}^{N} (e_k - \bar{e})^2}{N - 1} \tag{16}$$

9.2.6 *Step 6*—Analyze the measurement results, e_k,$\bar{e}$, and s^2, to determine the performance of the SUT with regard to the vendor's specifications pertaining to the user's application per Section 10.

10. Performance Evaluation

10.1 This section is specifically for the application of this test method for performance evaluation pertaining to the user's application. Four performance limits are used in this test method for performance evaluation, and a statistical test is described for each in the following sections.

10.2 *Introduction:*

10.2.1 After the data has been collected as specified in Section 9 and the error associated with each data point calculated as described in Section 8, the results shall be evaluated. Performance evaluation takes the form of using statistical tests to verify whether the SUT is operating within the vendor's claimed performance limits.

10.2.2 A vendor's performance specification is verified if the performance tests in this standard accept the null hypothesis with a p-value of greater than 0.95. The analysis is described in statistical terms as a combination of null and alternative hypotheses, written as H_0 and H_a, respectively. In Table 1, the four statistical tests used in this test method are described in terms of the null and alternative hypotheses. For example, Test I, the Average Error Test, applies to the expected average error, $E[\bar{e}]$, and the vendor's specified performance limit. If H_0 is true in a statistical sense, then measurement results obtained from the SUT are expected to be less than the vendor's specified performance limit, δ_{avg}, so the performance specification is accepted. In this case, the SUT is referred to as being within the vendor's performance specifications. Alternatively, if H_a is true in a statistical sense, then measurement results obtained from the SUT are not expected to be less than the vendor's performance specification limit, δ_{avg}. In this case, the SUT is referred to as being outside of the vendor's performance specifications.

10.2.3 In Table 1, the four tests used in this test method are listed with their associated performance specifications. The vendor's performance specifications are:

δ_{avg} = The vendor's specified performance limit on the expected average error, $E[\bar{e}]$;
δ_{quan} = The upper bound on the vendor-specified pth quantile of the average error, q_p;
δ_{max} = The maximum average error, e_{max}; and
σ_0^2 = The vendor's specified performance limit on the variance of the average error, σ^2.

10.2.4 In X1.1, a more detailed explanation of performance acceptance/rejection with regard to Tests I and II is provided. In particular, for Test II, if the experiment were repeated many times, $100 \times p$ percent of the trials will be less than δ_{quan}. When $p = 0.5$, Test II is a statement about the median error. An explanation of how one can determine the appropriate test for a given application is given in X1.1.

10.3 *Evaluating Performance*—This section describes the procedure for determining if the performance of the SUT is within the vendor's specifications for δ_{avg}, δ_{quan}, δ_{max}, and σ_0^2. In the following subsections, the processes for determining whether the performance of the SUT is within the vendor's specifications based on Tests I through IV are summarized.

10.3.1 *Average Error Test:*

10.3.1.1 With the assumption that the measurement error is normally distributed (see Appendix X1), using the Z-test, the SUT is not within the vendor's performance specifications if the following is true:

$$\frac{\bar{e} - \delta_{avg}}{\sqrt{s^2/N}} > Z_\alpha \tag{17}$$

where:

$\bar{e}$ = average measurement error computed using Eq 15,
s^2 = sample variance defined in Eq 16, and
Z_α = value at which the cumulative distribution function for the standard normal distribution has the value 0.95 (see Ref **4**). Specifically $Z_\alpha = 1.6449$.

10.3.1.2 See X1.3 for a more detailed explanation of the value of the Z test.

10.3.2 *Quantile Error Test:*

10.3.2.1 Let T be equal to the number of elements of $\{e_1, ..., e_N\}$ for which $e_k \leq \delta_{quan}$ is true. Using the Sign Test, the performance of the SUT is not within the vendor's specifications if the following is true:

$$T \leq b_{N,\alpha} \tag{18}$$

TABLE 1 Statistical Tests for the Analysis of Pose Measurement Systems

Test	Test Name	Null Hypothesis	Alternative Hypothesis
I	Average error test	H_0: $E[e] \leq \delta_{avg}$	H_a: $E[e] > \delta_{avg}$
II	Quantile error test	H_0: $q_p \leq \delta_{quan}$	H_a: $q_p > \delta_{quan}$
III	Maximum permissible error test	H_0: $e_{max} \leq \delta_{max}$	H_a: $e_{max} > \delta_{max}$
IV	Precision error test	H_0: $\sigma^2 \leq \sigma_0^2$	H_a: $\sigma^2 > \sigma_0^2$