Hierarchical Characterization Procedures for Dimensional Metrology

MacKinnon, David; Beraldin, Jean-Angelo; Cournoyer, Luc; Carrier, Benjamin

Publisher's version / Version de l'éditeur:
https://doi.org/10.1117/12.872124
Society of Photo-Optical Instrumentation & Electronics & Society for Imaging Science and Technology - Electronic Imaging 2011 Conference, 2011

Hierarchical Characterization Procedures for Dimensional Metrology

David MacKinnon[†], Jean-Angelo Beraldin., Luc Cournoyer, and Benjamin Carrier
Institute for Information Technology, National Research Council Canada, Ottawa, ON, Canada

ABSTRACT

We present a series of dimensional metrology procedures for evaluating the geometrical performance of a 3D imaging system that have either been designed or modified from existing procedures to ensure, where possible, statistical traceability of each characteristic value from the certified reference surface to the certifying laboratory. Because there are currently no internationally-accepted standards for characterizing 3D imaging systems, these procedures have been designed to avoid using characteristic values provided by the vendors of 3D imaging systems. For this paper, we focus only on characteristics related to geometric surface properties, dividing them into surface form precision and surface fit trueness. These characteristics have been selected to be familiar to operators of 3D imaging systems that use Geometrical Dimensioning and Tolerancing (GD&T). The procedures for generating characteristic values would form the basis of either a volumetric or application-specific analysis of the characteristic profile of a 3D imaging system. We use a hierarchical approach in which each procedure builds on either certified reference values or previously-generated characteristic values. Starting from one of three classes of surface forms, we demonstrate how procedures for quantifying for flatness, roundness, angularity, diameter error, angle error, sphere-spacing error, and unidirectional and bidirectional plane-spacing error are built upon each other. We demonstrate how these procedures can be used as part of a process for characterizing the geometrical performance of a 3D imaging system.

Keywords: Dimensional metrology, 3D imaging systems, statistical methods, geometrical properties, GD&T

1 INTRODUCTION

Dimensional metrology is concerned with all aspects of spatial measurements. A variety of systems exist for generating surface models of structures; however, few standards exist to evaluate the geometrical performance of these systems. The NRCC-IIT 3D Imaging and Metrology Research Project was formed, along with other research projects, to work with other research institutes in Canada and around the world to participate in the process of defining and refining standards for 3D imaging systems. The nomenclature and variable naming convention has been selected to be familiar to operators of the 3D imaging System Under Test (SUT) who use GD&T procedures. This differs from the more common approach based on nomenclature used by those familiar with Computerized Measuring Machines (CMMs).

The NRCC-IIT 3D Imaging and Metrology Research Project has adopted a hierarchical approach to artifact-based characterization and verification[1] of the performance of 3D imaging systems. This approach does not depend on characteristic values published by the equipment manufacturer but can be used to compare the performance of 3D imaging systems. All testing procedures utilize reference surfaces that would be certified to have certain spatial characteristics, and each spatial characteristic has associated with it a measure of the uncertainty associated with that characteristic value. Statistical analyses are used to verify the performance of the SUT. Finally, testing procedures have been developed to minimize the number of depth maps generated by each scanner while maintaining statistical validity.

The procedures presented in this paper represent the minimum required to quantify each characteristic being tested. These procedures would then be used as part of either a volumetric or application-specific analysis that has been adapted to each type of 3D imaging system. Discussion of volumetric and application-specific analysis methods is beyond the scope of the current paper. The ISO 17025[17] recommends that test procedures should be selected to suit the application to which the SUT will be placed so not all tests presented here will be required for all systems. Discussion is also limited to procedures that quantify the geometrical performance of the SUT. Procedures to assess the optical properties, resolution, and model fidelity of the SUT are also under development but have not been included in this paper. It is assumed that all tests are performed under constant and consistent environmental conditions (temperature, humidity, ambient illumination, etc.) whose contribution to the uncertainty budget is small enough to be neglected.

[†] david.mackinnon@nrc-cnrc.gc.ca

1.1 Terminology

The terminology used in this report is consistent with the terminology defined in the VIM[1], GUM[2], and ASTM E 2544-8[4]; however, care must be taken when using characteristic values published for individual 3D imaging systems because not all published characteristics adhere to the definitions in the VIM[1], GUM[2], or ASTM E2544-8[4]. This section summarizes the ISO and ASTM definitions of the terms precision and trueness, and indicates how each is measured. The term accuracy is not discussed in this document because it is an abstract, so unquantifiable, statistical concept.

1.1.1 Precision

Precision, illustrated in Figure 1, represents the closeness of agreement, or dispersion, of measured values and is quantified using measurement repeatability, intermediate precision, and measurement reproducibility. Each term can be described as follows:

Repeatability: Repeatability is evaluated under test conditions that do not change[1][4]. Repeatability is evaluated by performing multiple repetitions of the test procedure.

Intermediate precision: If most test conditions, such as the location and reference material remain the same but conditions like operator, calibration, or test date change then the term intermediate precision is used[1][4]. Intermediate precision is evaluated by performing the same test procedure using a different operator, calibration, test date, or a combination of these.

Reproducibility: If the varied test conditions include different locations and test objects then the reproducibility of the system is being evaluated[1][4]. Reproducibility is evaluated in collaboration with other laboratories who repeat the test procedure using their own staff and artifacts.

If it can be assumed that the measured values follow a Gaussian distribution, then precision can be quantified as measurement uncertainty. Measurement uncertainty is represented in dimensional metrology by the standard measurement uncertainty u [1][2], which is represented by the experimental standard deviation of the mean[2].

A distinction can also be made for 3D imaging systems between image-level and test-level uncertainty. An image is composed of multiple measured results so the image-level uncertainty of these measured results compared to a surface model can be assessed. We use N to represent the number of measurement results within each image. Each image can also be considered a single "measurement" of a surface so multiple images must be acquired to generate test-level measurement uncertainty. Repeatability, intermediate precision and reproducibility are all quantified by test-level uncertainty, and we use K to represent the number of images.

1.1.2 Trueness

Trueness, illustrated in Figure 1, is the closeness of agreement between a set of measured results and a reference value[1], and is quantified by measurement bias[1][4]. Also known as the systematic measurement error, measurement bias is a consistent, or systematic, deviation of the average measured result from a reference value[1] so can only be quantified where a reference value exists.

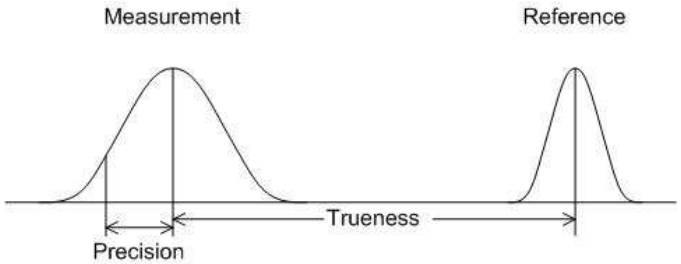


Figure. 1. Graphical representation of precision (dispersion of measurement results about the average) and trueness (disparity between the average measurement result and reference value).

1.1.3 Metrological versus Statistical Traceability

A characteristic value is considered to be metrologically traceable if an unbroken chain of calibrations, each with a stated uncertainty, exists between the indication and a primary standard[1][17]; although it is generally sufficient to establish a link to a national standard. Metrological traceability of the results of a measurement device, such as a 3D imaging

system, would be formally established by an accredited laboratory when it is calibrated[17] and is beyond the scope of this document. We instead refer to the statistical traceability of a characteristic value when it is possible to establish the chain of uncertainties. To be metrologically traceable, procedures capable to generating statistically traceable characteristic values would need to be performed in adherence to ISO 17025; a process that involves documenting the uncertainties, documenting the measurement procedure, performance of all procedures by an accredited laboratory, establishing the metrological chain to the SI, and documentation of the intervals at which calibrations of the SUT are performed[1].

Metrological traceability of a measurement instrument applies to two procedures: calibration and validation[17]. Calibration is the process of quantifying the relationship between a measurement results and one or more Certified Reference Value (CRVs), which can then be used to apply a correction to the measurement result to minimize the bias. Validation is the process of verifying that the measuring instrument is adequate for a given purpose, where verification is the provision of evidence that a set of specified requirements have been met[1]. Where it is not practical to perform formal calibration we use the term optimization to refer to the process of identifying and correcting for biases. An optimized SUT is considered to be calibrated only if the optimization process is metrologically traceable.

2 HIERARCHICAL TESTING

One of the primary areas of concern in dimensional metrology is characterizing how well a set of N measured results $P \in \Re^{N\times 3}$ fit a Certified Reference Surface (CRS). The set of measured results is then used to create a best-fitting surface model (BFSM) $M \subset \Re^3$ where M is defined as a connected, orientable, two-dimensional manifold[9]. The precision of the SUT is evaluated by first examining the measurement dispersion after creating M with minimal constraints on the dimensions of M. The trueness of the SUT is then evaluated by quantifying the bias relative to the CRV provided with each CRS.

2.1 Form Precision

Surface fit precision is quantified by determining the image-level variation of P about M using the appropriate surface model class. Frequently, the surface model class is either a sphere or a plane, but other custom forms, referred to as freeform surfaces, can be used for specialized applications; however, the surface model class must be sufficiently defined so that P can be assumed to be normally distributed about M. Form is the lowest-order classification of surface topography, the other two being surface waviness (2^{nd} order) and surface roughness (3^{rd} to 5^{th} order)[22][23]. Discussion about the characterization of waviness and roughness is beyond the scope of this paper. The amount of image-level variation of P about M is represented in two ways: form spread and form uncertainty. Form spread is used to represent the size of the region within which most measurements about M should lie. Form uncertainty is the image-level uncertainty associated with the image defined by P.

The precision characteristic value associated with the CRS must be verified to be much smaller than the precision characteristic value associated with the SUT. For example, the VDI 2634 Part 2[5] recommends in some tests that the precision characteristic value associated with the CRS be at least 5 times less than the associate characteristic value generated for the SUT. This approach, however, ignores the effects of measurement uncertainty in both the test method and the instrument or method used to generate the precision characteristic value associated with the CRS. In this document it is recommended that the precision characteristic value associated with the CRS must also be statistically verified to be significantly smaller than the characteristic value generated by the SUT for the CRS to be considered valid.

Characterizing precision is highly dependent on the best-fit algorithm used to create M so it is important that only well-tested and widely-accepted fitting algorithms are used. Where possible, we use the same algorithms the National Institute of Standards and Technology (NIST) used as reference methods in their Algorithm Testing System (ATS)[11].

2.2 Fit Trueness

Once an appropriate surface model class has been identified, the trueness of the 3D imaging system can be evaluated for that surface model class. Trueness is evaluated by determining whether the bias between a characteristic value and the appropriate CRV is small enough to be non-significant. Each CRV is, in fact, a pair of values: one represents the reference value and the other represents the measurement uncertainty associated with the reference value.

2.3 Surface Response

In dimensional metrology, it is also important to characterize how the SUT responds to changes in surface features such as surface orientation, depth, curvature, edges and variations in surface reflectivity. The effect of surface changes is quantified using a combination of detection limits[2] and measurement uncertainty. Surface response characteristic values can be traceable or untraceable depending how they are calculated. Discussion about surface response characteristics is beyond the scope of this paper.

3 SURFACE FORM PRECISION

A common approach to surface fitting involves minimizing the L_2, or Euclidean, distance between P and M. The L_2 distance is found by

$$\left\|\vec{p}_i - \hat{p}_i\right\|_2 = \sqrt{\left(\vec{p}_i - \hat{p}_i\right)\left(\vec{p}_i - \hat{p}_i\right)^T} \tag{1}$$

where $\vec{p}_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix} \in P$ is an $N \times 3$ vector representing a measured point in the 3-dimensional data space, N is the number of measurement results, and $\hat{p}_i \in M$ is the corresponding expected measured value in M. The error model can be expressed as

$$\vec{p}_i = \hat{p}_i + \vec{\xi}_i \tag{2}$$

where $\vec{\xi}_i \in \Re^{1\times 3}$ is a normally-distributed random error vector with zero mean and covariance $S_{\vec{p}}$. The form of $S_{\vec{p}}$ depends on the technology behind the 3D imaging system under evaluation.

The surface forms can be divided into simple linear forms, simple nonlinear forms, and complex nonlinear forms. Simple linear forms, also known as planes, can be fit using a closed-form solution so are the simplest class of surface form. Simple nonlinear forms are solved iteratively, with the speed of solution depending on how close the initial guess is to the final solution, and includes any other surface for which a derivative can be provided. Complex nonlinear forms consist of all other surfaces that require an iterative search to find a BFSM but are too complex to easily calculate a derivative to speed the search. Complex nonlinear forms can be fit using iterative closest point (ICP) procedures[18][19][20] but are not discussed in this paper.

3.1 Simple Linear Forms: Planar Surfaces

Let P be a data set obtained by scanning a planar CRS and in which all edge-affected regions have been removed. Because P should fit a planar-class surface model, M can be obtained using a procedure referred to as orthogonal least-squares regression[6][7][8] or Total Least-squares Regression (TLR)[8][10]. An example of a CRS for which M can be assumed to be a planar-class surface model can be seen in Figure 2(a). If the residuals are normally distributed about M then the approximation is also a maximum likelihood estimator (MLE), provided the measurement uncertainties for all axes are identically distributed[8].

The TLR solution can be obtained using a variety of approaches but the simplest and most widely accepted is Singular Value Decomposition (SVD). The SVD approach is applicable only if the measurement uncertainty of the 3D imaging system can be assumed to be primarily along each of the three spatial axes, are statistically independent of each other, and the uncertainty is approximately the same along each axis. Figure 2(b) shows that these assumptions may not be valid in some cases so more research is required into more robust plane-fitting methods. SVD was the reference algorithm for planar fitting used by National Institute of Standards and Technology (NIST) in their Algorithm Testing System (ATS)[11]. The TLR solution generates the normal $\vec{n}$ of M from the eigenvector associated with the smallest eigenvalue of the SVD solution. Combined with the centroid $\bar{p} = \begin{bmatrix} \bar{x} & \bar{y} & \bar{z} \end{bmatrix}$ of P, the planar BFSM can be uniquely and completely described.

3.1.1 Plane-fit Residual Uncertainty

If the assumption that the residuals are normally-distributed holds true then precision to a planar CRS can be quantified using the Type A standard form uncertainty for a planar model, referred to here as the plane-fit residual uncertainty

u_{plane}. We first find the signed orthogonal-fit residuals $\delta_i = \vec{n} \bullet (\vec{p}_i - \overline{p})$, then calculate the standard error of the estimate (SEE) for the planar BFSM using

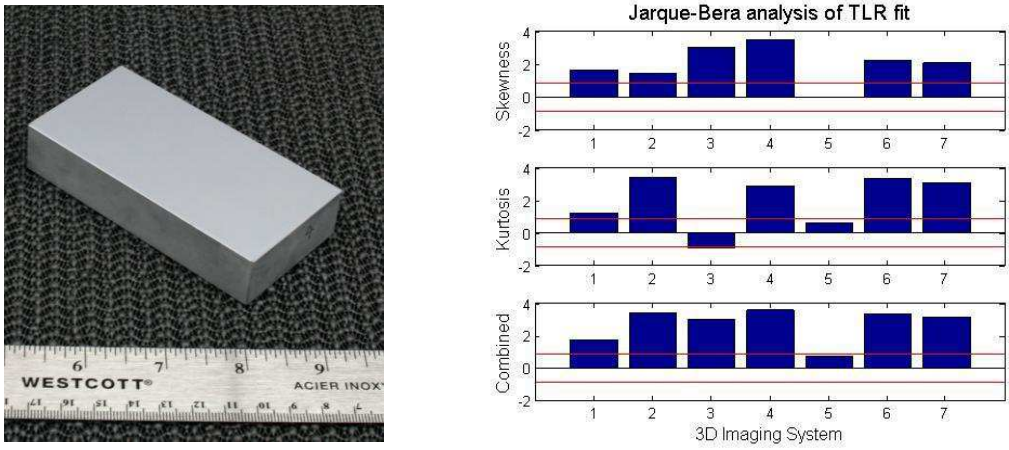
$$s_{plane} = \sqrt{\frac{\sum_{i=1}^{N} \delta_i^2}{N - \nu_{\dim}}} \tag{3}$$

where N is the number of measurements in P [3] and $\nu_{\dim}$ is the number of dimensions[2][3], in this case $\nu_{\dim} = 3$. Equation (3) represents the experimental standard deviation[2] of the measurements about (orthogonal to) the planar BFSM[3]. The plane-fit residual uncertainty is then obtained using

$$u_{plane} = \frac{s_{plane}}{\sqrt{N}} \tag{4}$$

and is an unbiased estimator of the population standard deviation[2][12]. Equation (4) somewhat overestimates the true planar measurement uncertainty because u_{plane} is a combination of the surface roughness and surface waviness of the CRS, and the measurement uncertainty of the SUT. The planar measurement uncertainty is an example of image-level uncertainty.

The concept of the plane-fit residual uncertainty is often unfamiliar to those who approach the assessment of 3D imaging systems from the perspective of GD&T. For this reason, we do not include it in to the list of geometrical properties used to assess the SUT. The plane-fit residual uncertainty is not itself sufficient to indicate the quality of the SUT unless it is provided with respect to one or more invariant reference quantities such as depth, lateral position, surface angle, surface reflectivity, etc., such as would be provided within a volumetric or application-specific analysis. Moreover, u_{plane} is only valid if the best-fit plane was generated using TLR, limiting its applicability to data sets in which the residuals can be assumed to be normally-distributed. As can be seen in Figure 2(b), this may not always be a valid assumption. Several less restrictive approaches to planar fitting are being explored including generalized TLR-ICP[20], renormalization[25], and least-median squares regression[16][26].



(a) (b)

Figure. 2. Flatness. (a) Diffusely-reflecting planar reference surface for assessing Flatness. (b) Jarque-Bera statistics for skewness, kurtosis and normality obtained from 7 triangulation-based laser range scanners. Bars that extend outside the horizontal bounding lines deviate significantly from the assumptions required for TLR fitting. The top graph represents the assumption of symmetry, the middle the assumption of peakedness, and the bottom is the combined assumption of fitting a normal distribution.

3.1.2 Flatness

People who are accustomed to working with GD&T quantify the precision of a planar surface using the Flatness metric[21]. According to the ASME Y14.5, a surface is considered flat if all measurements lie within a tolerance region bounded by planes parallel to the best-fit median plane[16]; however, the TLR plane is generally easier to obtain for large

data sets so is used instead of the median plane. Similarly, the precision with which the SUT is able to measure a planar surface can be represented by the Flatness characteristic F. We define F as the signed minimum distance between $\vec{p}_i$ and the BFSM

$$F = \max_{P_{trimmed}} \{\delta_i\} - \min_{P_{trimmed}} \{\delta_i\} \tag{5}$$

where $P_{trimmed}$ is a subset of P after no more than 3 measurement results per thousand with the largest absolute δ_i have been removed. Flatness, therefore, represents the ability of the SUT to determine whether a planar surface destined for a particular application meets tolerance limits. Equation (5) is equivalent to the flatness measurement error defined in the VDI 2634 Part 2[5] and is a composite of the uncertainty introduced by the SUT and the roughness and waviness of the surface of the CRS. Unlike u_{plane}, F is a non-parametric measure of precision so can be used with any plane-fitting method.

If the residuals are normally distributed with negligible surface roughness and waviness then $F \approx 6s_{plane}$; however, an examination of 7 triangulation-laser-based 3D imaging systems, shown in Figure 2(b), indicates that this assumption should be treated with caution. Jarque-Bera (JB) tests[14][15] of skewness (top graph), kurtosis (middle graph) and overall normality (bottom graph) indicate that most of the scanners tested deviate significantly from normality, as indicated by bars extending outside the horizontal bounding lines. As a result, it should not be assumed that $6s_{plane}$ can be used to represent Flatness for all 3D imaging systems.

Given a set $\{F_0,\ldots,F_K\}$ of F values generated as part of a volumetric or application-specific analysis, the F value associated with the SUT is $F = \max\{F_0,\ldots,F_K\} = \|F_j\|_\infty$, also known as the L_∞ norm of the set. It represents the maximum tolerance limits that can be assessed by the SUT under the conditions of the analytical protocol so indicates the performance of the SUT under worst-case operating conditions. It is, therefore, important to assess the applicability of the analytical protocol to the expected operating conditions of the SUT before drawing conclusions about the probable performance of the SUT.

3.1.3 Flatness Repeatability versus Reference

The F characteristic associated with a single scan has little meaning without information about the repeatability of F so multiple data sets P_j must be acquired under repeatability conditions to obtain the uncertainty u_F associated with the F value. Moreover, for the CRS to be considered a useful reference, the certified Flatness of the CRS $F_{ref} \pm u_{ref}$ must be significantly smaller than $\overline{F} \pm u_F$ where

$$\overline{F} = \frac{\sum_{i=0}^{K} F_i}{K} \tag{6}$$

is the average F value and

$$u_F = \sqrt{\frac{\sum_{i=0}^{K} \left(F_i - \overline{F}\right)^2}{K(K-1)}} \tag{7}$$

is the repeatability of F. Unlike u_{plane}, u_F is an example of test-level uncertainty. The significance of the difference between $\overline{F}$ and F_{ref} can be evaluated using

$$H_0: \quad \frac{\overline{F} - F_{ref}}{\sqrt{u_F^2 + u_{ref}^2}} \le z_\alpha \tag{8}$$

where H_0 denotes the null hypothesis and z_α is the z-statistic at a significance level α for a single-tailed test. Typically $\alpha = 0.05$ is used so $z_\alpha \approx 1.65$. If (8) is false then $\overline{F}$ is significantly larger than F_{ref} and the CRS is a valid reference. In this case, $\overline{F} \pm u_F$ is provided as the Flatness characteristic of the SUT under repeatability conditions to represent the expected range of F values that could be generated for a given planar surface.

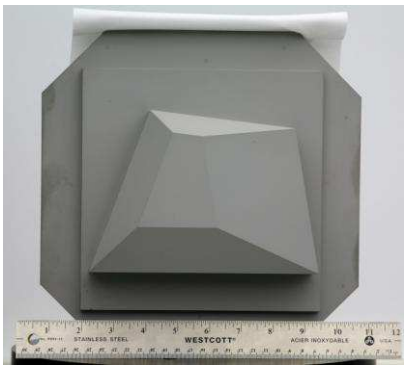
It should be noted that z_α best approximates the sample distribution if many images ($K > 30$) were used to generate u_F. If few images were used to generate u_F then $t_{\alpha,K-1}$ is a more appropriate statistic. The z-statistic for a given α is smaller than the equivalent t-statistic so (8) would fail more often if z_α is used. Everywhere that the z-statistic is used in this document, the t-statistic should be substituted when the number of images used to generate the repeatability statistic of the characteristic value is 30 or less.

3.1.4 Angularity

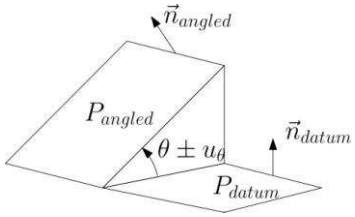
In GD&T, Angularity A is the Flatness of a planar surface P_{angled} under the constraint that it is at a specified angle θ to a reference, or datum, plane P_{datum}. A datum is a surface feature that serves as a spatial reference for a characteristic value so P_{datum} is defined as a plane that serves as a reference planar surface on the CRS. Angularity, Perpendicularity, and Parallelism are classes of Orientation tolerances, where A is the Orientation in which the angle in radians between $\vec{n}_{angled}$, the normal of P_{angled}, and $\vec{n}_{datum}$, the normal of P_{datum}, is $0 < \theta < \pi/2$.

For 3D imaging systems, A represents the effect of surface orientation on the F value generated by the SUT. When performing volumetric or application-specific analysis, P_{datum} will typically be fixed to some spatial position and orientation so that the u_{plane} of the reference is fixed. The A value associated with the SUT is the worst-case F value generated using a volumetric or application-specific analysis in which only θ is varied with respect to P_{datum}. Given a set $\{F_0, \ldots, F_K\}$ of F values in which only θ is varied, the A value associated with the SUT is $A = \|F_j\|_\infty$.

Figure 3(a) shows a CRS that can be used to assess the Angularity of the SUT. The CRS consists of a series of planar surfaces oriented at different angles to the central planar surface. The CRS can also be used to assess the Angle Error of the SUT, discussed in Section 4.2, and the Unidirectional Plane-spacing Error, discussed in Section 4.3.



(a)



(b)

Figure. 3. Angularity and Angle Error. (a) Reference surface for assessing Angularity, Angle Error, and Plane-spacing Error. (b) Illustration of how Angle Error is determined.

3.2 Simple Nonlinear Forms: Spherical Surfaces

Let P be a data set obtained by scanning a spherical CRS and in which all edge-affected regions have been removed. Because P should fit a spherical-class surface model, M can be generated using a nonlinear least squares fit to a sphere. In this case, the construction of M is constrained only in that the distance from all points on M to the center of the sphere are the same. An example of a CRS for which M can be assumed to be a spherical-class surface model can

be seen in Figure 4(a). The CRS in Figure 4(a) can also be used to assess Diameter Error, discussed in Section 4.1, and Sphere-spacing Error, discussed in Section 4.2.

3.2.1 Unconstrained Sphere Fitting

Characterizing sphere fit precision requires performing sphere fitting with the radius unconstrained. There is no closed form solution so this approach requires an iterative procedure initialized with an estimate of the location of the sphere centre $\hat{p}_c = \begin{bmatrix} \hat{x}_c & \hat{y}_c & \hat{z}_c \end{bmatrix}$ and the sphere radius estimate $\hat{r}$. The objective is then to seek a sphere center $\vec{p}_c = \begin{bmatrix} x_c & y_c & z_c \end{bmatrix}$ such that the objective function

$$f = \sum_{i=1}^{N} \left(\left\| \vec{p}_i - \vec{p}_c \right\|_2 - r \right)^2 \tag{9}$$

is minimized. The speed of convergence to an optimal solution depends on both how close the initial guess of $\vec{p}_c$ is to the solution and on the search algorithm used to seek the solution. Several variants of the search algorithm exist but the most popular is the Levenberg-Marquardt (LM) method. The LM method is considered to be the standard method for solving nonlinear least-squares problems[13] and is used by NIST in their ATS[11].

3.2.2 Sphere-fit Residual Uncertainty

If the assumption that the residuals are normally-distributed holds true then precision to a spherical CRS is can be quantified using the Type A standard form uncertainty for an unconstrained spherical model, referred to here as the sphere-fit residual uncertainty u_{sphere}. Using the signed orthogonal-fit residuals $\delta_i = \left\| \vec{p}_i - \hat{p}_c \right\|_2 - \hat{r}$, we calculate the SEE for the spherical BFSM as

$$s_{sphere} = \sqrt{\frac{\sum_{i=1}^{N} \delta_i^2}{N - \nu_{\dim}}} \tag{10}$$

where $\nu_{\dim} = 3$. Equation (10) represents the experimental standard deviation of the measurements orthogonal to the surface of the spherical BFSM. The sphere-fit residual uncertainty is obtained using

$$u_{sphere} = \frac{s_{sphere}}{\sqrt{N}} \tag{11}$$

which somewhat overestimates the true spherical measurement uncertainty because u_{sphere} is a combination of the roughness and waviness of the surface of the CRS, and the measurement uncertainty of the SUT.

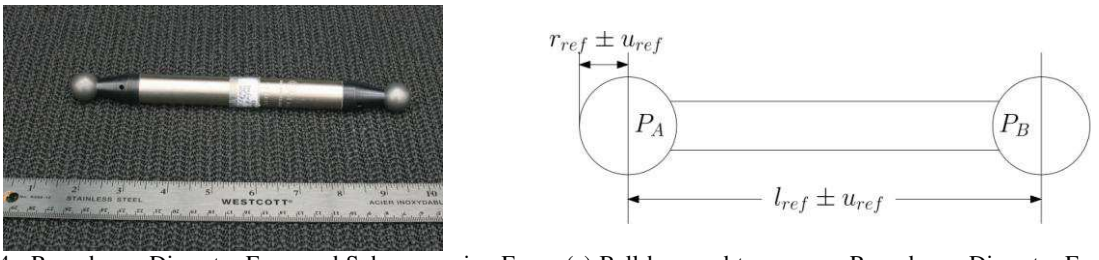


Figure. 4. Roundness, Diameter Error and Sphere-spacing Error. (a) Ball-bar used to measure Roundness, Diameter Error, and Sphere-spacing Error. (b) Illustration of how to measure Sphere-spacing Error.

3.2.3 Roundness

The ASME Y14.5[21] defines Roundness to represent the width of the region orthogonal to the surface of the sphere bounded by two spheres within which all measurements are found. Similarly, we define the Roundness characteristic R as representing the region orthogonal to the surface of the spherical BFSM within which at least 99.7% of all measurements should be found. The signed minimum distance between $\vec{p}_i$ and the BFSM, R is defined as

$$R = \max_{P_{trimmed}} \left\{ \delta_i \right\} - \min_{P_{trimmed}} \left\{ \delta_i \right\} \tag{12}$$

where $P_{trimmed}$ is a subset of P after no more than 3 measurement results per thousand with the largest absolute δ_i have been removed. Equation (12) is equivalent to the probing error defined in the VDI 2634 Part 2[5] and the VDI 2617 Part 6.2[24] but overestimates the true Roundness of the SUT because it is a composite of the uncertainty introduced by the SUT, and the roughness and waviness of the surface of the CRS. If the residuals are normally distributed then $R \approx 6s_{sphere}$.

3.2.4 Roundness Repeatability versus Reference

For the CRS to be considered a useful reference, the certified Roundness of the CRS $R_{ref} \pm u_{ref}$ must be significantly smaller than $\overline{R} \pm u_R$ where

$$\overline{R} = \frac{\sum_{i=0}^{K} R_i}{K} \tag{13}$$

is the average R value and

$$u_R = \sqrt{\frac{\sum_{i=0}^{K} \left(R_i - \overline{R}\right)^2}{K(K-1)}} \tag{14}$$

is the repeatability of R. . The significance of the difference between $\overline{R}$ and R_{ref} can be evaluated using the method described in Section 3.1.3. If the null hypothesis is false then $\overline{R}$ is significantly larger than R_{ref} and the CRS is a valid reference.

4 SURFACE FIT TRUENESS

When CRVs are provided with a CRS, form trueness can be evaluated for the 3D imaging system. Each CRV consists of a pair of numbers, one indicating the reference value and the second indicating the Type A standard uncertainty associated with the reference value. In this section, we present diameter error, angle error, sphere-spacing error, and plane-spacing error.

4.1 Diameter Error

Characterizing the trueness of a sphere diameter estimate involves seeking only an estimate of the sphere center location that minimizes the Euclidean distance and, by extension, the spread of diameter values associated with each measurement result. The process is similar to the unconstrained radius case described in Section 3.2.1 except that $\hat{r}$ is not obtained during the search process. The CRS shown in Figure 4(a) is an example of a surface that can be used to evaluate the sphere diameter trueness of a 3D imaging system.

The sphere form spread is not a useful metric in the constrained radius case because the residuals are rarely normally distributed about r_{ref}. Instead, we define the diameter error to be

$$E_D = 2\left(\frac{\left(\sum_{i=1}^{N} \left\| \vec{p}_i - \hat{p}_c \right\|_2\right)}{N} - r_{ref}\right) \tag{15}$$

where r_{ref} is the CRS sphere radius. A positive E_D indicates that the measured diameter is larger than $2r_{ref}$. The diameter error uncertainty is found using

$$u_D = 2\sqrt{u_{sphere}^2 + u_{ref}^2} \tag{16}$$

where u_{sphere} is the sphere-fit residual uncertainty obtained using (11) and u_{ref} is the uncertainty associated with the sphere radius CRV. Equation (16) arises from the law of propagation of uncertainties[2] and forms part of the statistical traceability chain[1] from E_D through the CRS to the certifying laboratory.

The significance of the difference between the measured and certified diameters can be evaluated using

$$H_0: \left|\frac{E_D}{u_D}\right| \le z_{\alpha/2} \tag{17}$$

where H_0 denotes the null hypothesis and $z_{\alpha/2}$ is the z-statistic at a significance level of $\alpha/2$ for a two-tailed test. Typically $\alpha = 0.05$ is used so $z_{\alpha/2} \approx 1.96$. If (17) is true then the measured sphere diameter is not significantly different than the certified sphere diameter.

4.2 Sphere-spacing Error

An extension of the known radius case is quantifying how accurately the 3D imaging system can measure the separation between sphere centers. An example of a CRS that can be used to evaluate the sphere-to-sphere distance trueness of a 3D imaging system can be seen in Figure 4(a). Spheres are often used as reference points for localizing surfaces so the accurate measurement of the size of a structure depends on the ability to accurately and precisely measure the distance between sets of reference spheres.

The measured sphere-to-sphere separation l is obtained for multiple data sets P_j under repeatability conditions to obtain the test-level uncertainty u_l associated with l. We combined the data sets to obtain $\bar{l} \pm u_l$ where

$$\bar{l} = \frac{\sum_{i=0}^{K} l_i}{K} \tag{18}$$

is the average l value and

$$u_l = \sqrt{\frac{\sum_{i=0}^{K} \left(l_i - \bar{l}\right)^2}{K(K-1)}} \tag{19}$$

is the repeatability of l. We define the sphere-spacing error to be

$$E_{SS} = \bar{l} - l_{ref} \tag{20}$$

where l_{ref} is the CRS sphere-to-sphere distance. The spheres-spacing error is analogous to the sphere-spacing error as defined in the VDI 2634 Part 2[5] with the difference that the VDI requires only a single data set to be generated. It is also similar to the error of indication of the sphere-to-sphere length of a ball-bar as described in VDI 2617 Part 6.2[24]. Neither document suggests using statistical testing to establish whether $\bar{l} \pm u_l$ is statistically distinguishable from $l_{ref} \pm u_{ref}$.

Related to E_{SS}, which quantifies the trueness of the sphere-to-sphere distance, is the sphere-spacing uncertainty u_{SS}, which quantifies the precision of the sphere-to-sphere fit. The sphere-spacing uncertainty is found using

$$u_{SS} = \sqrt{u_l^2 + u_{ref}^2} \tag{21}$$

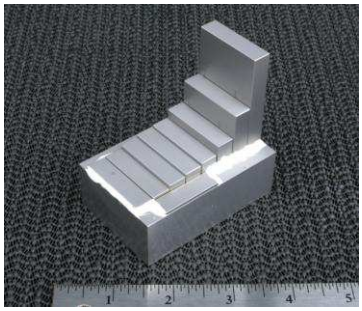
where u_{ref} is the uncertainty associated with the sphere-spacing CRV. The significance of the difference between the measured and certified sphere-to-sphere distance then takes the form

$$H_0: \left|\frac{E_{SS}}{u_{SS}}\right| \leq z_{\alpha/2} \quad . \tag{22}$$

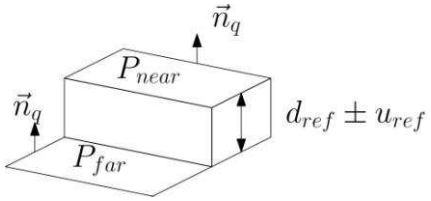
If (22) is true then the measured sphere-to-sphere distance is not significantly different than l_{ref}.

4.3 Unidirectional and Bidirectional Plane-spacing Error

Planes are often used to evaluated characteristics such as depth or range resolution; however, the ability of the 3D imaging system to represent the distance between parallel planes must first be assessed. An example of a CRS used to evaluate unidirectional plane-spacing error is shown in Figure 5(a). In practice, two planes are rarely coplanar so the BFSM is modified so that both planes have the same normal vector. This approach assumes that the planes are mostly coplanar and any non-coplanarity becomes part of the residuals.



(a)



(b)

Figure. 5. Plane-spacing Error. (a) Example of certified reference surface for Plane-spacing Error using stacked gauge blocks. (b) Illustration of how Plane-spacing error is determined.

Given two planes P_{near} and P_{far}, we can find the centroids $\bar{p}_{near}$ and $\bar{p}_{far}$ and use them to create a composite measured space

$$Q = \begin{bmatrix} P_{near} - \bar{p}_{near} \\ P_{far} - \bar{p}_{far} \end{bmatrix} \tag{23}$$

in which each plane has been translated to the origin independently. The composite normal $\vec{n}_q$, illustrated in Figure 5(b), is then generated using the TLR procedure described in Section 3.1. The distance between the planes along $\vec{n}_q$ becomes

$$d = \left| \frac{\vec{n}_q \cdot \left(\bar{p}_{near} - \bar{p}_{far}\right)}{\left\|\vec{n}_q\right\|_2} \right| \tag{24}$$

where $\left\|\vec{n}_q\right\|_2 = 1$ if $\vec{n}_q$ is a unit vector.

Multiple data sets P_j must be acquired under repeatability conditions to obtain the test-level uncertainty u_d associated with d. We combined the data sets to obtain $\bar{d} \pm u_d$ where

$$\bar{d} = \frac{\sum_{i=0}^{K} d_i}{K} \tag{25}$$

is the average d value and

$$u_d = \sqrt{\frac{\sum_{i=0}^{K}\left(d_i - \bar{d}\right)^2}{K(K-1)}} \tag{26}$$

is the repeatability of d.

The unidirectional plane-spacing error is obtained using

$$E_{UPS} = \bar{d} - d_{ref} \tag{27}$$

where d_{ref} is the CRV for the plane-to-plane spacing, illustrated in Figure 5(b). The unidirectional plane-spacing uncertainty is found using

$$u_{UPS} = \sqrt{u_d^2 + u_{ref}^2} \tag{28}$$

where u_{ref} is the uncertainty associated with d_{ref}. The significance of the difference between the measured and certified plane-to-plane distance can be evaluated using

$$H_0: \quad \left|\frac{E_{UPS}}{u_{UPS}}\right| \leq z_{\alpha/2} \quad . \tag{29}$$

If (29) is true then $\bar{d}$ is not significantly different than d_{ref}.

Bidirectional plane-spacing error E_{BPS} is similar to E_{UPS} except that it applies to 3D imaging systems, such as those employing a rotating bed, that can image planer surfaces with opposite surface normals. For these systems, planes with the same and opposite normals must both be considered. Equations (27) to (29) are applied to E_{BPS} and its associated uncertainty u_{BPS} the same way they were applied to E_{UPS} and u_{UPS}. Bidirectional plane-spacing error is similar to the error of indication of the length of a gauge block as described in VDI 2617 Part 6.2[24]; however, the VDI does not suggest using statistical testing to establish whether $E_{BPS} \pm u_{BPS}$ is statistically distinguishable from $l_{ref} \pm u_{ref}$.

4.4 Angle Error

Angle error E_a characterizes the ability of the SUT to represent the angular difference between two planar surfaces. Figure 3(a) shows an example of a CRS that can be used to assess E_a, Consider the two-plane system introduced in Section 3.1.4 in which the CRV of the difference in orientation between $\vec{n}_{angled}$ and $\vec{n}_{datum}$ is $\theta_{ref} \pm u_{ref}$, as illustrated in Figure 3(b). Unlike the previously-presented trueness metrics, the normal of a plane is a test-level variable so the repeatability of the surface normal must be generated from a set $\{P_0, \ldots, P_K\}$ of images. For each image, a pair of normals $\left\{\vec{n}_{angled}, \vec{n}_{datum}\right\}_i$ is generated and the angular difference, or inter-plane angle, between them is found using

$$\theta_i = \cos^{-1}\left(\frac{\vec{n}_{angled} \cdot \vec{n}_{datum}}{\left\|\vec{n}_{angled}\right\|_2 \left\|\vec{n}_{datum}\right\|_2}\right) \tag{30}$$

from which $\bar{\theta} \pm u_\theta$ is calculated as the test-level mean inter-plane angle and associated uncertainty. The angle error is then found using

$$E_a = \bar{\theta} - \theta_{ref} \tag{31}$$

to represent the signed different between the reference and measured angles. The angle uncertainty is found using

$$u_a = \sqrt{u_\theta^2 + u_{ref}^2} \tag{32}$$

so that the significance of the difference between the measured and certified inter-plane angle can be evaluated using

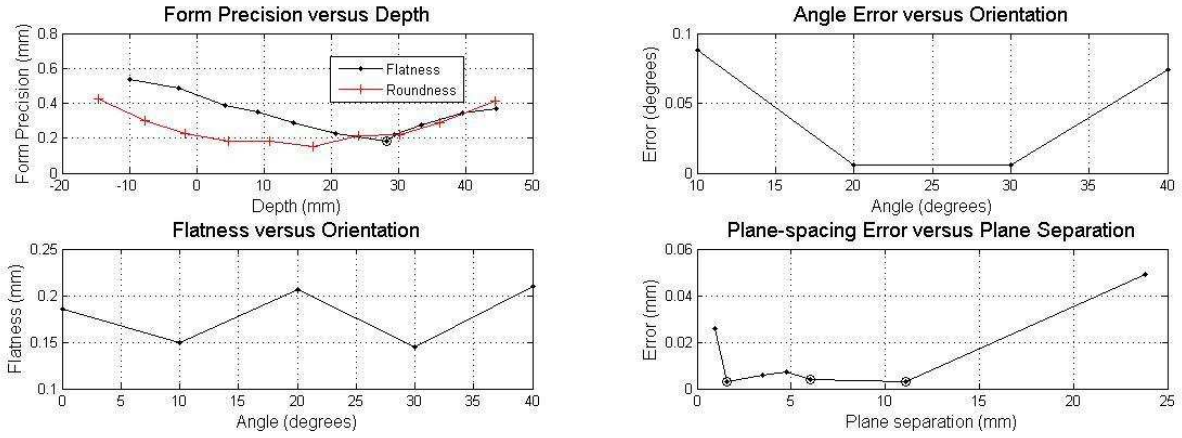
$$H_0: \left|\frac{E_a}{u_a}\right| \le z_{\alpha/2} \quad . \tag{33}$$

If (33) is true then the average inter-plane angle is not significantly different than the reference value.

5 EXAMPLE

We illustrate the application of these characteristics with a simple example. We use a short-range (less than 1 meter) laser line scanner as the SUT and perform a limited assessment of the performance of the SUT. All tests of significance were at a probability of less than 0.05. The test procedure is as follows:

- Plot F as a function of depth z within the expected operational depth-of-field DoF using the CRS shown in Figure 2(b) in which the planar surface normal is oriented along the z axis and the CRS is centered in the field of view. Three images are obtained at regular intervals through the DoF.
- Repeatability of F is assessed near the point of minimum F based on 10 images.
- Unidirectional Plane-spacing Error versus reference plane near the point of minimum F based on 10 images using the CRS shown in Figure 5(a).
- The CRS used to assess F is also used to assess A and E_a.
- Plot R as a function of z within the expected operational DoF using the CRS shown in Figure 4(a) in which the CRS is centered in the field of view. Three images are obtained at regular intervals through the DoF.
- Repeatability of R, E_d, and E_{SS} is assessed near the point of minimum R based on 10 images.



(a) (b)

Figure. 6. Test results . (a) Top graph: Flatness (dot markers) and Roundness (cross markers) versus Depth. Bottom graph: Flatness versus Orientation taken at a point indicated in top graph with a circle. (b) Top graph: Angle Error versus Orientation. Bottom graph: Unidirectional Plane-spacing Error versus plane-to-plane separation with points not significantly different than the reference value (prob<0.05) indicated with circles.

The top graph of Figure (6)(a) shows the average F value (solid line with dot markers) obtained through the operational DoF of < 50 mm (closest to the SUT) to $>$ -20 mm (farthest from the SUT). Flatness varied from a minimum of 0.186 ± 0.00387 mm at $z = 28.26$ mm to a maximum of 0.538 ± 0.00472 cm at $z = -9.85$ mm. In all cases, F values were significantly larger than 5 times the Flatness CRV so the CRS is considered to be sufficient to evaluate the Flatness of the SUT with an acceptable margin of error.

Angularity (solid line with cross markers), shown in the bottom graph of Figure (6)(a) , was assessed as the maximum F at $z = 28.26$ cm and was found to be 0.210 ± 0.00526 mm from a surface with orientation 40 degrees. All F values used to assess A were significantly larger than 5 times their respective CRVs so the CRS was considered to be sufficient to evaluate the Angularity of the SUT with an acceptable margin of error. Angle errors, shown in the top graph of Figure (6)(b) were significantly large for all orientations so recalibration may be required.

The top graph of Figure (6)(a) shows the average R value obtained through the operational DoF. Roundness varied from a minimum of 0.149 ± 0.00500 mm at $z = 17.41$ mm to a maximum of 0.427 ± 0.00260 cm at $z = -14.56$ mm. No Roundness CRV was provided for the CRS so the sufficiency of the CRS to evaluating the Roundness of the SUT could not be established. Diameter Error was assessed at all 10 depth positions, but in all cases was significantly large so recalibration would be required.

Sphere-spacing Error was assessed at all 10 points used to assess R, but in all cases was significantly large so recalibration would be required. Unidirectional Plane-spacing Error was assessed by placing the reference surface near the point of minimum F can calculating the error for 7 plane separation values. Figure (6)(b) shows E_{UPS} as a function plane separation with points in which the measured plane separation was not significantly different than the applicable CRV indicated with circles.

In summary, the CRS shown in Figure 2(b) was sufficient to assess F of the SUT, which varied from 0.186 to 0.538 within the operational DoF. The CRS shown in Figure 4(a) was sufficient to assess the R of the SUT, which varied from 0.149 to 0.427. Sphere diameter, sphere-to-sphere distance, and plane orientation were found to be significantly different than the applicable CRV so the SUT may require recalibration. Plane-to-plane distance was found to different significantly than the applicable CRV in more than half the tested plane separations so the SUT may require recalibration.

6 ADDITIONAL WORK

This paper covered only a portion of the test suit designed for use with the Metrology Testkit. The complete test suite includes measures of surface geometric, model fidelity, resolution, optical properties, and external frame of reference (EFOR). Model fidelity refers to how closely a measurement model of the SUT matches the reference model (plane, sphere or freeform surface) of the CRS. Resolution includes characteristics such as depth, structural, and intensity resolution. Optical properties focus on the ability of the SUT to respond to different optical properties of the surface and include, but is not limited to, characteristics such as dynamic range and E_{UPS} as a function of surface reflectivity. EFOR applies when the SUT uses an external frame of reference when generating a measurement model and includes systems in which the motion of the data acquisition unit must be tracked.

The test suit also includes procedures to measure SUT repeatability, intermediate precision, and reproducibility. These procedures can be adapted for a general volumetric analysis of the SUT or tailored for application-specific analysis. Also not covered in this paper were procedures for characterizing the geometric properties of localization and profile. Localization refers to the ability of the SUT to localize spatial structures with respect to a reference structure. Profile refers to the ability of the SUT to reproduce a freeform surface.

7 CONCLUSION

We have presented a series of dimensional metrology procedures designed to evaluate the geometrical performance of a 3D imaging system, specifically flatness, angularity, roundness, diameter error, sphere-spacing error, unidirectional and bidirectional plane-spacing error, and angle error. These characteristics, divided into surface form precision and surface fit trueness, are statistically traceability through a certified reference surface to a certifying laboratory. The procedures were hierarchically organized so that the suitability of the reference surface could be assessed prior to performing error analysis. Statistical methods were presented to first test whether flatness and roundness metrics are significantly larger than the equivalent for the certified reference surface, then whether diameter, sphere-spacing, unidirectional and bidirectional plane-spacing errors, and angle errors are not significantly different than zero.

REFERENCES

[1] JCGM 200:2008, "International Vocabulary of metrology – Basic and general concepts and associated terms (VIM)," JCGM/WG-2 (2008).

[2] ISO Guide 98-3, "Uncertainty of Measurement – Part 3: Guide to the Expression of Uncertainty in Measurement (GUM: 1995)," International Standards Organization (ISO) (1995).

[3] S. A. Glantz and B. K. Slinker, [Primer of Applied Regression and Analysis of Variance], Mcgraw-Hill, Inc. 23, (1990).

[4] ASTM E 2544-08, "Standard Terminology for Three-Dimensional (3D) Imaging Systems," ASTM, (2008).

[5] VDI/VDE 2634 Part 2, "Optical 3-D measuring systems - Part 2: Optical systems based on area scanning," Beuth Verlag GmbH (2002).

[6] G.C. Calafiore, "Outliers Robustness in Multivariate Orthogonal Regression," Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans 330(6), 674-679 (2000).

[7] J. P. Brooks and E. L. Boone, "Robust L_1 orthogonal regression," Technical Report, Virginia Commonwealth University (2008).

[8] I. Markovsky and S. van Huffel, "Overview of total least-squares methods," Signal Processing 87, 2283-2302 (2007).

[9] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface Reconstruction from Unorganized Points," Computers and Graphics 26, 71-78 (1992).

[10] C. C. Paige and M. Wei, "Analysis of the generalized total least squares $AX \approx B$ problem when some columns of A are free of error," Numerische Mathematik 65, 177-202 (1993).

[11] C. M. Shakarji, "Least-Squares Fitting Algorithms of the NIST Algorithm Testing System," Journal of Research of the National Institute of Standards and Technology 103(6), 633-641 (1998)

[12] W. Mendenhall, R. L. Scheaffer, and D. D. Wackerly, [Mathematical Statistics with Applications],. 3rd edition, Duxbury press, Boston, 303 (1986).

[13] [Numerical Recipes in C: The Art of Scientific Computing], Cambridge University Press, 683-685 (1992).

[14] G. Brys, M. Hubert, and A. Struyf, "A Robustification of the Jarque-Bera Test of Normality," in Proceedings of the CompStat 2004 Symposium (2004).

[15] D. Öztuna, A.H. Elhan, and E. Tüccar, "Investigation of Four Different Normality Tests in Terms of Type 1 Error Rate and Power under Different Distributions," Turkish Journal of Medical Science 36(3), 171-176 (2006).

[16] L.A. Sarabia, M.C. Ortiz, and X. Tomás, "Performance of the orthogonal least median squares regression," Analytica Chimica Acta 348, 11-18 (1997).

[17] ISO/IEC 17025:2005, "General requirements for the competence of testing and calibration laboratories," International Standards Organization (ISO) (2005).

[18] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," Image and Vision Computing 10, 145-155 (1992).

[19] P. Besl and H. McKay, "A method for registration of 3-D shapes," IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2), 239-259 (1992).

[20] R. S.-J. Estépar, A. Brun, and C.-F. Westin, "Robust Generalized Total Least Squares Iterative Closest Point Registration," Proceedings of the Seventh International Conference on Medical Image Computing and Computer-Assisted Intervention (2004).

[21] ASME Y14.5.1:1994, "Mathematical Definition of Dimensioning and Tolerancing Principles," ASME (1994).

[22] J. Raja, B. Muralikrishnan, and S. Fu, "Recent advances in separation of roughness, waviness and form," Precision Engineering 26(2), 222-235 (2002).

[23] M. Dietzsch, M. Gerlach, and S. Gröger, "Back to the envelope system with morphological operations for the evaluation of surfaces," Wear 264(5-6), 411-415(2008).

[24] VDI/VDE 2617 Part 6.2, "Accuracy of coordinate measuring machines: Characteristics and their testing - Guideline for the application of DIN EN ISO 10360 to coordinate measuring machines with optical distance sensors," Beuth Verlag GmbH (2005).

[25] Y. Kanazawa and K. Kanatani, "Reliability of Fitting a Plane to Range Data," IEICE Transactions on Information and Systems E78-D(12), 1630-1635 (1995).

[26] D. Mintz, P. Meer, and A. Rosenfeld, "Analysis of the least median of squares estimator for computer vision applications," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 621-623 (1992)