Progress in the specification of optical instruments for the measurement of surface form and texture

Peter J. de Groot
Zygo Corporation, Middlefield, CT 06457

ABSTRACT

Specifications for confocal microscopes, optical interferometers and other methods of measuring areal surface topography can be confusing and misleading. The emerging ISO 25178 standards, together with the established international vocabulary of metrology, provide a foundation for improved specifications for 3D surface metrology instrumentation. The approach in this paper links instrument specifications to metrological characteristics that can influence a measurement, using consistent definitions of terms, and reference to verification procedures.

Keywords: Calibration, Metrology, Specifications, ISO, VIM, GUM, Standards, Interferometry

1. INTRODUCTION

Just how accurate is your 3D measuring microscope? Is it better than competitive systems or technologies? How can you be confident that the measurements will detect part defects? How do you know if your metrology instrumentation is working properly? Proper specifications for manufactured instruments are in principle useful in answering these questions. Unfortunately, a variety of vague and inconsistent terms populates specification sheets, complicating instrument comparison and applications development.

Table 1. Published specifications related to instrument noise for a selection of commercially manufactured 3D interference microscopes. These instruments have comparable noise performance but divergent and inconsistent specifications.

Instrument	Specification	Value
A	Repeatability of surface RMS (Z)	0.003 nm
B	RMS repeatability (RMSσ)	< 0.01 nm
C	Vertical resolution	0.01 nm
D	RMS repeatability of surface accuracy	0.01 nm
E	RMS repeatability	<0.02 nm
F	Noise floor	0.05 nm
G	Vertical resolution	< 0.1 nm
H	Vertical resolution	0.1 nm
I	RMS repeatability	0.3 nm RMS
J	Vertical resolution	1 nm

Table 1 provides a typical example of the need for progress in specifying 3D topography instruments. The tabulated specifications and values are from published specification sheets for commercially manufactured optical coherence scanning interferometry (CSI) microscopes. The unifying concept behind all of these specifications is the intrinsic instrument noise present in the surface height map—a metrological characteristic that limits how much the reported surface topography varies from measurement to measurement. The diversity in the nomenclature is matched only by the divergence in the quantitative values, which ranges from 0.003 nm to 1 nm, more than two orders of magnitude. Although instruments do vary meaningfully in performance and capability, it is reasonable to assert that the largest contributor to the variations in Table 1 is the ambiguity in the meaning of the specification.

Quoted values for resolving power are no better. Table 2 shows a surprisingly wide range of values for lateral resolution, again for technologies and optical configurations that are not so different from each other, at least not as different as the factor of ten divergence would suggest.

Table 2. Published specifications related to lateral resolution for a selection of commercially manufactured 3D interference microscopes having similar resolving power. The notes are from the original specification sheets.

Instrument	Specification	Notes	Value
A	Optical resolution (X, Y)	surface dependent	0.4 – 0.6 µm
B	Lateral resolution	objective dependent	0.36 to 9.50 µm
C	Lateral resolution	Sparrow criterion	0.38 µm
D	Optical resolution white (L&S)	half diffraction limit	0.26 µm
E	Lateral resolution	[optional capability]	0.13 µm
F	Lateral resolution	150×	0.05 µm

These and other examples illustrate the confusing, contradictory and perhaps unintentionally misleading specifications that are common currency, even as precision metrology advances at an admirable pace. The problem is not simply a commercial one: the academic literature also requires a better way to report the capability of new technologies.

Many instrument makers and national metrology institutes (NMIs) are working with standards organizations to refine terminology, identify procedures for calibration and verification, and provide a foundation for meaningful performance specifications. This paper is a report of progress in this effort.

2. THE ROLE OF THE STANDARDS

2.1 Standardized metrology vocabulary

Recognizing the need to standardize terminology and methods for evaluating measurement uncertainty, the Joint Committee for Guides in Metrology (JCGM) has generated and maintains two key documents: *The International vocabulary of metrology* or VIM [1] and the *Guide to the expression of uncertainty in measurement* or GUM [2]. The GUM incorporates most of the VIM and is, therefore, the one reference document that every metrologist should have on hand.

Most recently, the International Organization for Standardization (ISO) has been developing the 25178 series of standards related to areal surface topography measurement, in part to assist in interpreting the GUM for these types of instruments. Some of the most recent work has been in the development of specification standards for well-known 3D topography-measuring tools such as confocal, stylus, focus variation and interference microscopy, encompassing terminology, calibration, and verification. Several of these standards have been published, while others are still in development. An example is ISO 25178-604:2013, abbreviated in this paper as ISO part 604, which refers to CSI [3]. Additional examples include part 601 for stylus instruments [4] and draft part 600 for general terms and definitions common to all areal surface topography instrumentation [5]. Although the ISO 25178 standards do not directly address the problem of instrument specification, they can serve as a foundation for uniform performance characterization and verification.

2.2 Identifying the "measurand"

Some specifications are relatively simple: height, width, weight, range of motion for stages, working distance for objectives and so on. Performance specifications related to accuracy and precision are another matter. Many performance quotes relate to a measurement capability, such as measurement of surface texture with a quoted level of precision or accuracy. Consequently, performance specifications should identify the *measurand*, defined as the quantity for which we are establishing a value by a measurement (GUM, B.2.9 [2]). A good example is the step height between neighboring surface areas. An extension of this concept is an array of measurands, as in the overall surface topography, which is the surface height as a function of surface position indexed by (x, y) coordinates or camera pixel location.

Although measurand definition may seem like a simple issue, this is where many instrument specifications go wrong. For example, many vendors quantify instrument noise by measuring the repeatability of the RMS roughness parameter *Sq*, perhaps with some averaging. This can easily be misinterpreted as the RMS noise for the individual height values on the surface map, which it is not. Instrument specifications often report the repeatability of the RMS parameter, not because it is informative, but because it is an impressively small number.

Another important issue is the role of lateral filtering. If, for example, we elect to apply a smoothing filter to the data, this reduces the random noise between image pixels, but it also changes the measurand. Instead of measuring the surface topography at the highest possible lateral sampling, we are measuring something else—the smoothed surface topography.

2.3 Sample types and measurement conditions

The ISO specification standards emphasize that the formal definition of the measurand should include any conditions of measurement that relate uniquely to the sample, as in the example: "the step height between surfaces regions A and B at 20 °C." For 3D measuring microscopes, particularly tools that can measure a wide variety of surface textures, we also need to make it clear that the specifications are confirmed using material measurement standards as defined in ISO part 70 that are optically smooth and, compatible with our measurement technology [6]. The instrument specifications ignore significant contributors to real measurements such as the instrument transfer function, variations in surface reflectivity, surface slope, thermal drifts, and distortions related to the measurement principle or adjustment problems. This is widely understood ("spec sheets are always under ideal conditions with pristine samples"), but it is worth stating it clearly for the record.

2.4 Metrological characteristics

Table 3. Performance specification mapping.

Metrological characteristic	Instrument specifications	Notes
Amplification coefficient (z)	· Step height repeatability · Step height accuracy	Expressed in height units and/or as a percentage of the measured height
Linearity deviation (z)	· Height response linearity	Expressed as a maximum permissible error (MPE)
Measurement noise	· Surface topography repeatability · Repeatability of the RMS	Expressed as a standard deviation for each specifications
Topographic lateral resolution	· Optical lateral resolution · Lateral sampling	The specifications are for influence factors that *relate* to lateral resolution
Residual flatness	(not specified)	Calibrated and adjusted in situ using a *system error subtract* procedure
Field amplification and linearity (xy)	(not specified)	Calibrated and adjusted *in situ*

The ISO 25178 series defines metrological characteristics that may influence the results of determining a quantitative value for the measurand [3]. Table 3 below lists in the first column the metrological characteristics deemed relevant by

ISO TC 213 WG 16 to areal surface topography measurement methods. An approach to developing sensible instrument specifications is to associate these metrological characteristics with the expected uncertainty contribution of a specific measurement type, as illustrated by the second column of Table 3. Importantly, these specifications include all relevant data processing and measurement conditions with sufficient detail to allow for verification.

A manufacturer's performance specifications should allow us to estimate what the uncertainty of the most basic measurement tasks using the quoted value from the manufacturer (a "type B" uncertainty analysis according to NIST and to the GUM [2, 7]). The uncertainty for a measurement will necessarily include other factors. Nonetheless, the instrument specifications serve as a baseline guide in determining how best to configure an application, or even whether the task is feasible with the instrument.

In what follows, I discuss a few example specifications consistent with this approach.

3. INSTRUMENT SPECIFICATIONS

3.1 Surface topography repeatability

The ISO standards recognize *instrument noise* (part 604, 2.1.9 [3]) as the internal noise added to the output signal caused by the instrument if ideally placed in an externally noise-free environment. More generally, there is also *measurement noise* (part 604: 2.1.10 [3]), defined as noise added to the output signal occurring during the normal use of the instrument. Measurement noise includes the instrument noise as well as external contributors arising from the environment (thermal, vibration, air turbulence) and other sources.

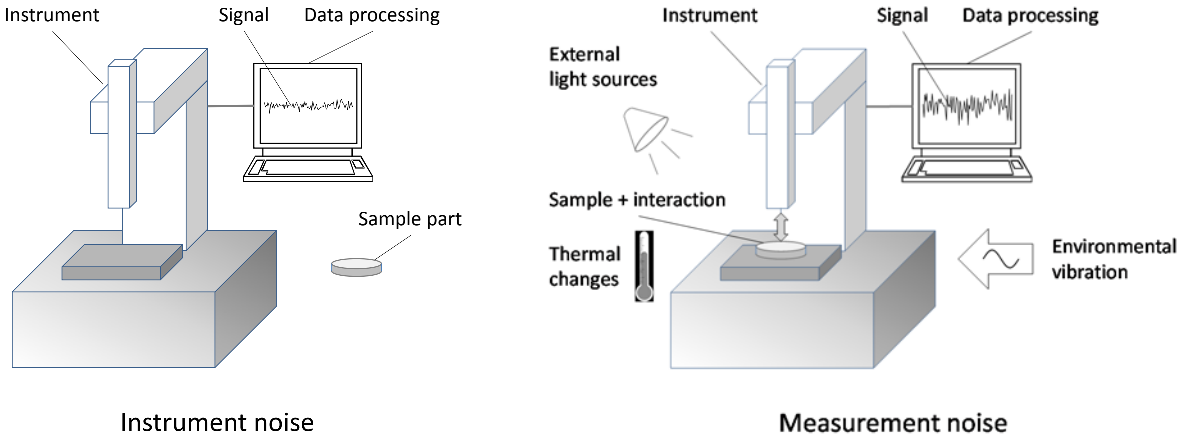

Figure 1: Illustration of the distinction between instrument noise (left) and measurement noise (right). Most specifications refer to instrument noise, which is an idealization of the measurement assuming no external disturbances (adapted from part 600 [5]).

A basic test of measurement noise consists of measuring twice to see if a measurement is reasonably consistent with repeated trials. If it is, then we can make adjustments that allow us to make future measurements with confidence that the results have an acceptable level of error with respect to the true value. For 3D measurement, the most fundamental output is an areal surface topography map, which shows the height over an array of perhaps a million different surface points corresponding to camera pixels. The *surface topography repeatability* (part 604, 2.1.11 [3]), tells us how close we can expect the indicated height value for a specific sample point (or camera pixel) to repeat if we measure it over and over again without changing the conditions of measurement (Figure 2). The value is a root-mean-square (RMS) or standard deviation, and is readily computed from statistics over a full image of surface points.

An evaluation of the repeatability under idealized conditions, with a sample considered to be compatible with the measurement principle, provides a measure of the intrinsic instrument noise that is often quoted as a basic specification for commercial instruments. For optical systems based on specular surface reflection, a suitable sample is a durable flat part with a mirror-like finish, such as a silicon carbide (SiC) flat. The simplest repeatability test is to measure a suitable flatness artefact twice, subtract the difference of the two resulting topography maps, and calculate the standard deviation *Sq* of the difference map [8, 9]. The surface topography repeatability, denoted here as the ISO measurement noise N_{M}, is then

$$N_{\mathrm{M}} = \frac{\mathrm{Sq}\left(\mathit{difference}\right)}{\sqrt{2}} . \tag{1}$$

For situations where intermittent disturbances such as vibrations may influence the repeatability, or where a more authoritative value is essential, a more reliable statistical approach is to take a large number of repeated measurements of the artefact. Several methods treat the statistics of these multiple measurements to provide a robust estimate of repeatability [10, 11].

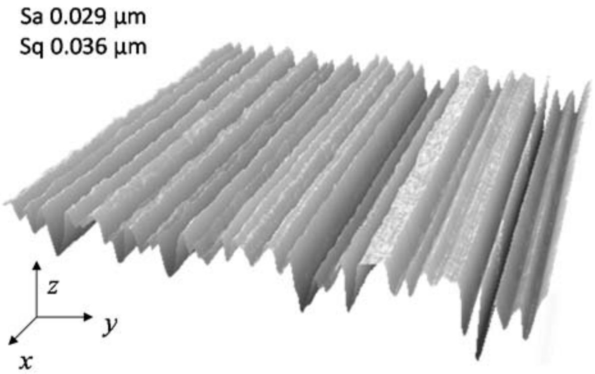

Figure 2: Surface topography repeatability relates most closely to the *z*-direction noise level for individual image points distributed over the *x,y* field. Other specified repeatability values may refer to surface texture parameters such as the *Sa* or *Sq* [12]. The distinction should be unambiguously clear, as should be the measurement time bandwidth as well as any relevant filtering.

Lateral smoothing, or the use of an S-filter as described in ISO 25178 part 3, improves the surface topography repeatability value [13, 14], and may even be mandated for correlation to other measurement techniques that may have intrinsically different spatial frequency responses [15]. For a formal specification, however, it is essential to state any lateral filtering operations, as these may have an influence on other specifications, such as lateral sampling.

From this discussion, we can conclude that surface topography repeatability is the ISO-recognized term most closely associated with the intrinsic instrument noise of the system, when evaluated under ideal conditions. So what about the other terms listed in Table 1, which we currently see frequently in commercial specifications?

The term *vertical resolution* in standardized metrology vocabulary refers in most cases to the number of digits displayed at the output (VIM 4.15 [1]; GUM F.2.2.1 [2]), rather than a fundamental performance limitation of the sensing technology. In many instruments, the only difference between *low* and *high* resolution is the digital storage allocation, for example 8 bits or 16 bits. Consequently, it is preferable not to use *vertical resolution* as a synonym for noise level or detectability of small surface heights, except as a qualitative term.

With respect to ambiguous term *RMS repeatability*, if the specification relates to a derived parameter such as the RMS of the surface topography, the correct way to describe the effect of noise would be to specify the *repeatability of the RMS parameter*, or even better, the *repeatability of the Sq parameter*.

Finally, given that specifications related to instrument repeatability are associated with random noise, it is essential to provide a corresponding time bandwidth for the specification. Surface topography repeatability improves with averaging or by using a data acquisition that takes more time and acquires more raw data samples prior to processing the result [11]. Indeed, given enough time, it is possible to achieve almost any quoted noise level, even in the picometer range. It is misleading, therefore, to compare the repeatability specifications of an instrument that measures at 10 Hz with another that measures at 0.1 Hz, without noting the difference in measurement time.

3.2 Step height accuracy

The determination of a discrete height difference between neighboring flat areas is one of the more common measurements for areal surface topography instruments. The measurand is the reported value of a step height standard (SHS), most commonly a purchased artefact or *material measure* having a structure similar to that shown in Figure 3. A comparison of the reported value to the certified value gives us an indication of how well the instrument can make measurements of discrete heights. Often one will see an accuracy specification associated with this evaluation.

The VIM discourages the use of the term *accuracy* as a quantitative attribute (VIM, 2.13, Note 1 [1]), declaring it a purely qualitative concept, as in "phase shifting interferometry is very accurate." I would argue, however, that accuracy can be a quantitative performance attribute if clearly associated with a well-defined measurement performance test. A definition that is close to the common understanding of step height accuracy, is the uncertainty in the measured step height after averaging repeated trials to remove random noise. This specification is most closely associated with *systematic errors* (VIM, 2.17 [1]) in the *amplification coefficient* (part 604, 2.1.8), which is the scaling factor that we use when reporting surface heights.

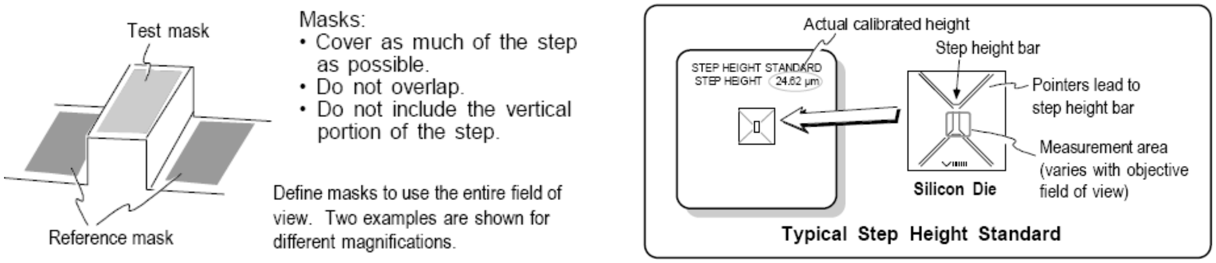

Figure 3: Example shape and measurement techniques for a step height standard (SHS) [16].

The usual method for validating this specification is to measure one or more SHSs bearing certification from a NMI or qualified laboratory [6, 17]. It is essential to recognize that an SHS itself has an intrinsic uncertainty, often greater than the quoted specification for the instrument. There is also an uncertainty associated with the measurement procedure, for which there is not yet a formal areal topography standard (the 2D profile standard is ISO 5436-1 [18]) although there are proposed procedures available [16, 19, 20]. For these reasons, it is not necessarily advisable to adjust the instrument to conform to the SHS value unless the reported value falls outside the uncertainty limits of the SHS. A perfect example of this is a HeNe 633 nm laser Fizeau interferometer, which scales height data to an independent realization of the meter [21, 22]. It would be senseless to adjust such an instrument to match exactly the stated value for a mechanical SHS, as this would only degrade the uncertainty. Likewise, many instruments include built-in calibrations using capacitance gages or even laser interferometers [23, 24] , potentially obviating the need to establish the amplitude coefficient using NMI-calibrated artefacts, although these artefacts are always of value to verify that the metrology system is in good working order.

3.3 Step height repeatability

A companion specification to the step height accuracy is the step height repeatability, which reports how well repeated step height measurements agree with each other. The measurement technique involves a great deal of averaging over many image pixels, so the influence of random pixel noise is small. Rather, this test simultaneously evaluates the repeatability of height scaling factor and the measurement procedure itself, which may involve automated area masking or data trimming. The uncertainty in individual step height measurements comprises the contributions from repeatability, the amplification coefficient and the linearity [19].

3.4 Height response linearity

Instrument manufacturers often characterize and, if possible, correct nonlinearity during manufacture, and the outcome is an acceptance test limit such as maximum permissible error. A candidate specification provides the largest deviation of the response curve with respect to a best-fit line, as illustrated in Figure 4. For instruments such as confocal and CSI microscopes that rely on the fidelity of a *z*-axis mechanical scanning device to measure surface heights, it is often sufficient to calibrate this mechanism, assuming no other significant source of nonlinearity.

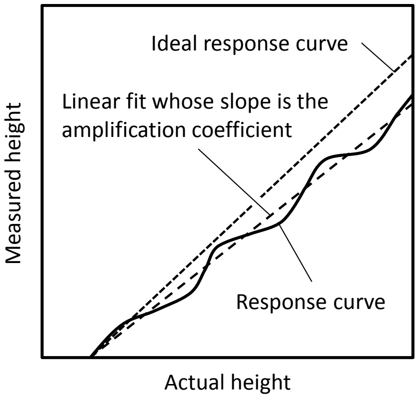

Figure 4: Height response curve, showing the actual response, the response corrected for nonlinearity, and the ideal linear response with the correct amplification coefficient (adapted from part 601 [4]).

Although common for stylus tools, a linearity specification is uncommon for optical surface topography measurement instruments. This is unfortunate, as the linearity of the height response of an interferometric microscope is an intrinsic property that is more difficult to calibrate and adjust in the field than the amplification coefficient, so a specification can be quite helpful. The manufacturer usually calibrates and adjusts for linearity during assembly using traceable length metrology such as a laser displacement interferometer. Verification feasible using a set of step heights of different size [19], measurements of a small step at various positions in the scan—a technique that simultaneously provides a better measure of the amplification coefficient than is possible with a single SHS.

In some instruments, additional hardware integrates a laser displacement gage into the mechanical scanning system of a CSI microscope [23-25], allowing for continuous, independent calibration of the amplification coefficient and correction for nonlinear response. This does not diminish the utility of a linearity specification, but it should reduce the quoted nonlinearity, and perhaps increase confidence in its stability over time.

3.5 Lateral resolution

Instrument response when evaluating surface topography depends on many factors, one if which is the proximity of neighboring features. For optical instruments such as interference and confocal microscopes, the ability to correctly report height differences typically declines monotonically as the separation of neighboring features becomes smaller, eventually reaching a point where the features are indistinguishable. In ISO 25178, the *topographic lateral resolution* is the metrological characteristic most closely associated with this phenomenon [3, 5].

There is some flexibility in quantifying and validating the topographic lateral resolution according to the intended application. A comprehensive characterization of height response as a function of spatial frequency is the *instrument transfer function* or ITF (part 604, 2.1.19 and C.1 [3]). The ITF idea parallels the traditional optical imaging concept of a modulation transfer function or MTF. The ITF is strongly influenced by hardware, software and adjustment factors, many of which are unstable with time. Figure 5 shows that with sufficient lateral sampling, the ITF and MTF are in fact in close agreement for small surface height deviations [26]. One way to summarize the ITF in a single number is the *lateral period limit* (part 604, 2.1.17) given by the spatial period at which the ITF falls to 50 %. Methods are available for evaluating the ITF *in situ* using artefacts such as step heights for determining the edge spread function [27, 28], or periodic structures, for directly visualizing the lateral period limit [29-31].

A more familiar concept of resolving power is the specification of the smallest center-to-center separation of features that still allows us to see clearly that there are two features present. Figure 6 shows on the right two closely-spaced trenches formed by patterning silicon on a quartz substrate [32]. The interference microscopy 3D image on the left-hand side of the figure shows that there are indeed two lines present, although they appear blurred at this high magnification. As the center-to-center separation between the lines decreases, the resolving power of the instrument becomes a limiting factor in determining if the two lines are clearly separated in the 3D image. If the resolving power is insufficient, the two lines appear as one larger line. A number of methods are available for evaluating lateral resolution experimentally using an artefact similar to that shown in Figure 6, with a selection of line pairs to establish the distance at which the line images merge [33].

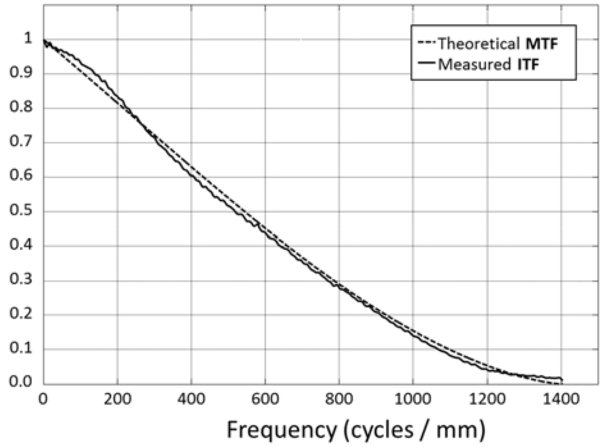

Figure 5: Comparison of the theoretical optical MTF and the measured magnitude of the 3D topography ITF for a CSI instrument at 20×. The vertical axis is the ratio of the reported height to the actual height of sinusoidal topography patterns at the indicated spatial frequency. The lateral period limit is at approximately 520 cycles/mm or 1.9 µm spatial period, while the calculated smallest resolvable feature separation is 0.67 µm, using the Sparrow criterion [26].

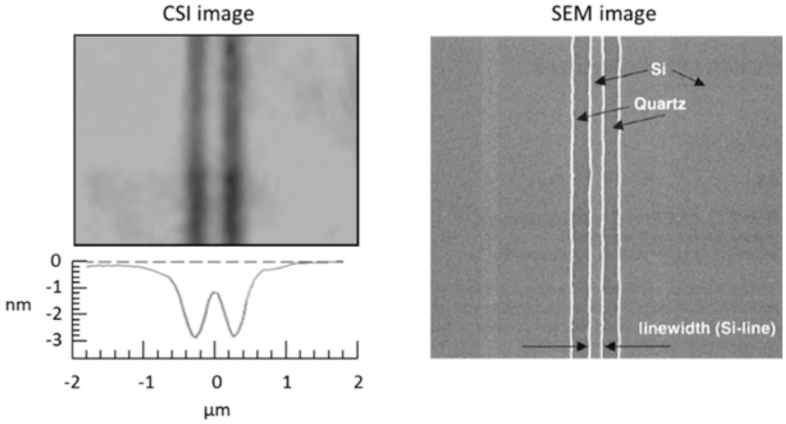

Figure 6: Left: 3D CSI microscope image of two parallel trenches of a standard from Supracon AG, using a 100× Mirau objective with an NA of 0.85. Right: Corresponding scanning electron microscope image, showing the linewidth and the sample structure. The linewidth is 200 nm and the center-to-center spacing of the lines is 440 nm [34].

An alternative approach to specifying lateral resolution is by quantifying the *influence factors* that have an impact on the instrument's ability to resolve surface topography features (part 604, 3.2 [3]). Several of these, including the wavelength λ and the numerical aperture A_N of the objective, group sensibly together to establish the *optical lateral resolution*, which refers uniquely to the properties of the optical system, irrespective of detection, data processing or interpretation methods (part 604, C.1 [3]). Similarly to the ITF, there is an argument that traditional intensity-imaging principles extend to optical lateral resolution, hence the frequent use of the Rayleigh and Sparrow criteria in specifications for surface topography instrumentation. For a diffraction-limited microscope with an incoherent light source imaged into the pupil of an objective having a numerical aperture A_N, the Sparrow limit (part 604, 2.3.8 [3]) of an imaging microscope entirely free of optical imperfections is

$$R_l = 0.47\,\lambda / A_N \quad . \tag{2}$$

Clearly this number does not tell the whole story; but it does give an idea of the ultimate limits on the resolving power of an optical instrument—additional factors, such as the camera format and data processing, can have a large effect on the ITF; but it is unlikely that surface features spaced closer than the Sparrow limit can be properly resolved.

An entirely different, separate influence factor is the *lateral sampling* or *sampling interval* in the (*x*,*y*) field (part 604, 2.1.12, 2.1.13 [3]). This value relates to the camera and available magnifications. The measurand here is the distance between two entirely independent signal samples on the object surface, sometimes referred to as the object-space pixel size. To take advantage of the diffraction-limited optical resolution, the camera pixel size in object space must be much smaller than R_l , otherwise the resolution is said to be camera limited.

This all seems straightforward, so how do we explain the divergence in quoted specification in Table 2? Part of the answer is the confusion between the topography lateral resolution, which quantifies the resolving power of the instrument as a whole, and the influence factors such as optical lateral resolution and lateral sampling. This is easily repaired by using more explicit terminology, such as *optical* lateral resolution or lateral *sampling*, recognizing that other factors come into play when measuring parts in practice. Another solution is to avoid non-standard measures of lateral resolution, such as minimum detectable linewidth, or half the diffraction limit. Finally, while post-processing to sharpen edges or otherwise adjust the ITF may enhance the appearance of small features; it also increases noise while not actually increasing the resolving power of the instrument.

Another pitfall is that the specification of lateral resolution is for a *capability*, not an error source; therefore, lateral resolution does not map directly to an uncertainty budget. The uncertainty in a measurement of the relative (*x*,*y*) locations of isolated features, for example, is only indirectly limited by the lateral resolution—with sufficient signal to noise, it is possible to locate isolated feature edges laterally in the field of view with a precision of a few nm, far below the quoted lateral resolution value. The effect of lateral resolution on the surface heights is quantified by the complex-valued ITF, a function that is absent from specification sheets given its complexity and variability.

4. CALIBRATION AND ADJUSTMENT

Some of the metrological characteristics listed in Table 3 are frequently unspecified, with the recommendation to determine their uncertainty contributions by means of *in situ* calibration and adjustment.

The field of view of an areal surface topography instrument comprises measurement points on an (*x*,*y*) grid, with height values in the *z* direction (part 604, 2.1.2 [3]). The reported (*x*,*y*)positions within a field of view are influenced primarily by optical effects for imaging microscopes and by stage mechanisms for point-scanning instruments. The most common calibration is for the field amplification coefficient in the (*x*,*y*) plane, which is directly related to magnification. A further refinement corrects for linearity and perpendicularity, using an interpolation method or the equivalent [35]. These calibrations and adjustments require a standard artefact, often an array of features [36].

Figure 7: Example cross-grating artefacts for calibration of the field amplification and linearity [17, 36, 37].

The idealised topography detected by the instrument when viewing a hypothetical perfect flat part is the *residual flatness*, which essentially is the height bias as a function of image position (part 604, 2.1.21 [3]). The simplest approach is to measure a standard artefact such as a high-quality silicon carbide reference flat [9]. The result is a *system error file* that can be subtracted from all subsequent measurements [38]. A more complete procedure requires a manual or automated sequence of multiple measurements with small lateral displacements of the artefact so as to average out the artefact roughness [39].

Of course, nothing inhibits a manufacturer from specifying these metrological characteristics or at a minimum, the influence factors that might limit the ability to calibrate them. For example, a potential specification is the reproducibility of the residual flatness calibration, or even the expected flatness without *in situ* calibration, using a factory adjustment. Similarly for the field amplification and linearity, some known limits on the uncalibrated optical distortion would be useful in determining if an *in situ* calibration is required for a specific metrology task.

5. EXAMPLE PERFORMANCE SPECIFICATIONS

Table 4. Example specifications for a CSI instrument [40]. A global footnote references relevant documents such as the CSI standard ISO 25178-604 and the material measures standard ISO 5436-1.

Specification	**Value**	**Footnotes**
Surface topography repeatability	0.1 nm	Repeatability (1σ)) for SmartPSI mode, 1 second data acquisition, 1 million image points, 3 × 3 pixel surface filtering
Repeatability of the RMS	0.005 nm	Repeatability (1σ) of the ISO Sq parameter, same conditions as for the surface topography repeatability
Optical lateral resolution	0.33 μm	100×, 0.85 NA objective, Sparrow criterion
Lateral sampling	0.04 μm	100× objective, 2× zoom
Step height accuracy	0.8 %	Instrument contribution to the uncertainty (k = 1) for step height measurement in extended scan mode (0.15 mm to 20 mm range)
Height response linearity	≤20 nm	Maximum deviation with respect to the best fit linear response
Step height repeatability	0.1 %	Repeatability (1σ) in reported step height as verified using a 1.8 μm and 24 μm standard artefacts

6. PRACTICAL METROLOGY

Modern metrology instruments for areal surface topography accommodate a wide variety of surface types. Real-world performance depends on surface texture, surface slope, thermal drifts, form distortions and other influence factors related to the interaction of the surface structure with the illumination. While specifications can summarize an instrument's basic capabilities, nothing replaces adjusting and measuring an actual part which the system is intended to measure in application, in the actual environment in which the system is intended to be used [2]. Quality control experts have understood this for some time, which is the reason for empirical tests such as the classic test for gage repeatability and reproducibility (GR&R), performed on the part of interest close to the actual conditions of measurement. These tests combine with correlation studies that relate the functional properties of the part to the detected surface structure.

7. CLOSING REMARKS

I chose the title of this paper carefully: This is a report on *progress*, not on a conclusion. Instrument providers and standards organizations are actively working to improve confidence in metrology, supplemented by international standards documents, good practice guides [41], and a growing body of literature. Instrument specifications are improving with the introduction of new products. However, there is (yet) no established, mandatory or standardized list of specifications. The idea is to provide a methodology while allowing for flexibility in the choice of which specifications and verification procedures are most relevant to user needs.

As a final note, specialized industries, such as semiconductor equipment, are developing standards appropriate to their own requirements. These standards in many cases incorporate generally useful concepts regarding reference metrology, uncertainty and gage capability [42].

REFERENCES

[1] JCGM, [200:2012 International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd Edition] Joint Committee for Guides in Metrology (2012).

[2] JCGM, [100:2008 Evaluation of measurement data—Guide to the expression of uncertainty in measurement (GUM)] Joint Committee for Guides in Metrology (2008).

[3] ISO, [25178-604:2013(E): Geometrical product specification (GPS) – Surface texture: Areal – Nominal characteristics of non-contact (coherence scanning interferometric microscopy) instruments] International Organization for Standardization, Geneva (2013).

[4] ISO, [25178-601: Geometrical product specifications (GPS) – Surface texture: Areal – Part 601: Nominal characteristics of contact (stylus) instruments] International Organization for Standardization, Geneva (2010).

[5] ISO, [WD 25178-600:2013(E): Geometrical product specifications (GPS) — Surface texture: Areal — Part 600: Metrological characteristics for areal-topography measuring methods] International Organization for Standardization, Geneva (2013).

[6] ISO, [FDIS 25178-70:2013(E): Geometrical product specification (GPS) -- Surface texture: Areal -- Part 70: Material measures] International Organization for Standardization (2013).

[7] Taylor, B. N., and Kuyatt, C. E., [NIST Technical Note 1297: Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results], National Institute of Standards and Technology, (1994).

[8] VDI/VDE, [2655 Part 1.1: Optical measurement and microtopographies; calibration of interference microscopes and depth setting standards for roughness measurement] Verein Deutscher Ingenieure/Verband der Elektrotechnik, Elektronik und Informationstechnik, Beuth-Verlag (2011).

[9] Giusca, C. L., Leach, R. K., Helary, F. *et al.*, “Calibration of the scales of areal surface topography-measuring instruments: part 1. Measurement noise and residual flatness,” Measurement Science and Technology 23(3), 035008 (2012).

[10] Haitjema, H., and Morel, M. A. A., “Noise bias removal in profile measurements,” Measurement 38(1), 21-29 (2005).

[11] Zygo Corporation, [Measuring Sub-Angstrom Surface Texture], Application note AN-0002 (2008).

[12] ISO, [25178-2: Geometrical product specifications (GPS) — Surface texture: Areal — Part 2: Terms, definitions and surface texture parameters] International Organization for Standardization, Geneva (2012).

[13] ISO, [25178-3:2012(E): Geometrical product specifications (GPS) — Surface texture: Areal — Part 3: Specification operators] International Organization for Standardization, Geneva (2012).

[14] Seewig, J. "Areal Filtering Methods," [Characterisation of Areal Surface Texture] Springer Berlin Heidelberg, 4 (2013).

[15] Boedecker, S., Bauer, W., Krüger-Sehm, R. *et al.*, "Comparability and uncertainty of shape measurements with white-light interferometers," Proc. SPIE 7718, 77180J-77180J-12.(2010).

[16] Zygo Corporation, [Step Height MetroPro Application], Operating manual OMP-0368E (2011).

[17] Leach, R. K., Giusca, C. L., and Rubert, P., "A single set of material measures for the calibration of areal surface topography measuring instruments: the NPL Areal Bento Box," Proc. Met. & Props, 406-413.(2013).

[18] ISO, [5436-1: Geometrical product specifications (GPS)-surface texture: profile method; measurement standards-part 1. Material measures] International Organization for Standardization, Geneva (2000).

[19] Giusca, C. L., Leach, R. K., and Helery, F., “Calibration of the scales of areal surface topography measuring instruments: part 2. Amplification, linearity and squareness,” Measurement Science and Technology 23(6), 065005 (2012).

[20] Leach, R., Brown, L., Jiang, X. *et al.*, [Guide to the Measurement of Smooth Surface Topography using Coherence Scanning Interferometry], National Physical Laboratory, Teddington (2008).

[21] Quinn, T. J., “Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001),” Metrologia 40(2), 103-133 (2003).

[22] Evans, C. J., and Davies, A. D., “Certification, self-calibration and uncertainty in optical surface testing,” International Journal of Precision Technology 3(4), 388 (2013).

[23] de Groot, P. J., Colonna de Lega, X., and Grigg, D. A., "Step height measurements using a combination of a laser displacement gage and a broadband interferometric surface profiler," Proc. SPIE 4778, 127-130.(2002).

[24] Olszak, A., and Schmit, J., “High-stability white-light interferometry with reference signal for real-time correction of scanning errors,” Optical Engineering 42(1), 54-59 (2003).

[25] Kiyono, S., Gao, W., Zhang, S. *et al.*, “Self-calibration of a scanning white light interference microscope,” Optical Engineering 39(10), 2720 (2000).
[26] Colonna de Lega, X., and de Groot, P., "Lateral resolution and instrument transfer function as criteria for selecting surface metrology instruments," OSA Proc. Optical Fabrication and Testing, OTu1D.(2012).
[27] de Groot, P. "Interpreting interferometric height measurements using the instrument transfer function," [FRINGE 2005] Springer Verlag, Berlin, Heidelberg (2006).
[28] Takacs, P. Z., Li, M. X., Furenlid, K. *et al.*, "Step-height standard for surface-profiler calibration," Proc. SPIE 1995, 235-244.(1993).
[29] Pehnelt, S., Osten, W., and Seewig, J., “Vergleichende Untersuchung optischer Oberflächenmessgeräte mit einem Chirp-Kalibriernormal,” Technisches Messen 78(10), 457-462 (2011).
[30] Giusca, C. L., and Leach, R. K., “Calibration of the scales of areal surface topography measuring instruments: part 3. Resolution,” Measurement Science and Technology 24(10), 105010 (2013).
[31] Weckenmann, A., Tan, Ö., Hoffmann, J. *et al.*, “Practice-oriented evaluation of lateral resolution for micro- and nanometre measurement techniques,” Measurement Science and Technology 20(6), 065103 (2009).
[32] Huebner, U., Morgenroth, W., Boucher, R. *et al.*, "Development of a nanoscale linewidth-standard for high-resolution optical microscopy," Proc. SPIE 5965, 59651W-59651W-9.(2005).
[33] Yu, G., and Tortonese, M., "Metrology standards for semiconductor manufacturing," Proc. 7th International Conference on Solid-State and Integrated Circuits Technology, 588-593 vol.1.(2004).
[34] de Groot, P., Colonna de Lega, X., Sykora, D. M. *et al.*, “The Meaning and Measure of Lateral Resolution for Surface Profiling Interferometers,” Optics and Photonics News 23(4), 10-13 (2012).
[35] Henning, A., Giusca, C., Forbes, A. *et al.*, “Correction for lateral distortion in coherence scanning interferometry,” CIRP Annals - Manufacturing Technology 62(1), 547-550 (2013).
[36] Zygo Corporation, [Precision Lateral Calibration Standard], Operating manual OMP-0484A (2003).
[37] Leach, R. K., and Giusca, C. L., "Determination of the metrological characteristics of optical surface topography measuring instruments," Proc. SPIE 8430, 84300Q-84300Q-7.(2012).
[38] Zygo Corporation, [NewView 7200 & 7300], Operating manual OMP-0536F (2011).
[39] Creath, K., and Wyant, J. C., “Absolute measurement of surface roughness,” Applied Optics 29(26), 3823 (1990).
[40] Zygo Corporation, [NexView Optical Profiler], Specification sheet SS-0095 09/12 (2013).
[41] Petzing, J., Coupland, J., and Leach, R., [The Measurement of Rough Surface Topography using Coherence Scanning Interferometry], National Physical Laboratory, Teddington (2010).
[42] Ukraintsev, V., and Banke, B., “Review of reference metrology for nanotechnology: significance, challenges, and solutions,” Journal of Micro/Nanolithography, MEMS, and MOEMS 11(1), 011010-1-011010-9 (2012).