Proc. of SPIE Vol. 1776, Interferometry: Surface Characterization and Testing, ed. K Creath, J E Greivenkamp (Dec 1992) Copyright SPIE

ISO figure specification (ISO 10110-5) and digital interferometry

Lars A. Selberg and Bruce E. Truax

Zygo Corporation
Middlefield, CT USA

Abstract

The ISO standard for optical drawing specification (ISO 10110) includes a part on surface form tolerances (figure error). Standardization of this specification provides a succinct and readily understood nomenclature which will be useful in industry, especially as it is intended to unify visual and digital evaluation. It is a peculiar standard in that it does not control surface form directly, but controls it by its effect in a particular test, interferometric comparison with a "perfect" reference. The application of this standard to digital interferometry is discussed, including a review of the mechanics of this standard and some examples. ISO 10110-5 is vague on some points and these are discussed. These issues include spatial frequency sampling, specific algorithms for calculating some of the tolerances, measurement accuracy concerns, aspherics and guidelines for determining if digital interferometry is required.

1. Introduction

Standards are essential to industry and commerce. In a global marketplace, international standards provide not only uniformity of specification, but language independent specification. The optics industry has had a variety of national standards (e.g., DIN 3140) and pseudo-national (MIL 141). At the time of this writing, the first international standard covering the specification of optical elements and assemblies ISO (International Standards Organization) 10110[1] is drafted and balloting is in process. The standard is comprised of thirteen parts, each covering some of the dimensions, specifications and tolerances which appear on a drawing. Part 5 of ISO 10110 describes the standard for specifying figure of optical surfaces, or "Surface Form Tolerances".

ISO 10110-5 takes on the difficult task of standardizing a measurement which is commonly made either by eye or with sophisticated measurement equipment (digital interferometers). The scope of this paper is to outline nomenclature use to specify figure and to discuss some of the implications this has for implementation through digital interferometry.

It is important to remember that standards are intended to clarify specifications, not necessarily simplify them. Thus, there are a few basic rules of thumb to observe. First, if using ISO 10110 to specify a part, make sure to state so explicitly on the drawing. Second, if the your requirements are not easily or adequately described by ISO 10110-5, then don't use it. Third, even if you use ISO 10110, it is still possible to use notes on the drawing to convey additional information.

2. Fundamental Assumptions

Surface form error (figure error) is a specification that controls a linear dimension, i.e., surface height variations from some nominal surface. Thus, one might expect that the standard for tolerancing surface form errors would be expressed in units of linear dimension. This would place the burden, if need be, on the metrologist to translate the specification into meaningful units for his or her test equipment. This is not the case. Surface form error is specified in terms of its effect in a specific type of test apparatus: double pass, normal incidence interference at 546 nm (Hg arc lamp). Therefore, the burden is on the metrologist to convert the specification if the test apparatus is something other than test plate and a green light box. While this is awkward for visual analysis with interferometers operating at other wavelengths, this generally means only the setting of a few controls.

3. Digital Analysis Setup Parameters

Before making a measurement with a digital interferometer it is necessary to set certain parameters so that the results will be scaled according to the specifications of this standard. The parameters listed below correspond to those found on the Zygo MARK IV*xp*, MARK IV, MARK III, and ZAPP systems. Other digital interferometer systems will have similar parameters.

Wavelength of the measurement (WAVE IN)

This is the wavelength of laser in the interferometer. For most interferometers this is 633 nm (HeNe laser). Some systems are available based on lasers with other wavelengths.

Wavelength of the output results (WAVE OUT)

This defaults to 546.07 nm by ISO 10110-5 and corresponds to the green emission line of mercury (Hg). This is the approximate wavelength of light boxes used with test plates. Check the drawing for a note which will specify a different output wavelength.

Units of the output

This should be selected for "fringes" or "waves". The choice here will affect the choice of scale factor.

Scale factor - waves to fringes (Interferogram Scale Factor, ISF)

If fringes are selected for the output units, the scale factor should be set to 0.5. If "waves" are selected for the output units, set the scale factor to 1.

Data sign

This is positive or negative, and determines if a feature on the surface is a bump or a hole. It has no effect on the calculated results specified by ISO 10110-5, but polishing maps will be wrong if this is not set correctly. This may be called "polarity".

4. Indication of the Specification

The specification is written as,

3 / A (B/C) RMSx < D (all ϕ E)

The quantities in this expression are calculated from the measured interferometric cavity data. As successive surfaces are fit to the data and subtracted, the PV and RMS of these surfaces and what they leave behind are the tolerances. The drawing indications specified by ISO 10110-5 are relatively straightforward.

A	Sagitta error	The tolerance on power/focus of the surface with respect to a reference sphere with the nominal RoC
B	Irregularity	The tolerance on the surface form error which remains after the sagitta error has been removed, i.e., the PV with respect to best fit reference sphere.
C	Rotationally symmetric irregularity (RSI)	The tolerance on the rotationally symmetric component of the surface form error after the best fit sphere has been subtracted, i.e., the PV of the Approximating Aspheric Surface (AAS, see definition below). This corresponds to edge roll, spherical aberration and/or zonal errors. In earlier drafts of the standard this was called "non-spherical, rotationally symmetric error", and the term still persists in a few sections of the standard at the time of this writing.

D	RMSx	This is the tolerance on any of the three versions of the RMS error. More than one of the RMSx may be specified. RMSt - Residual RMS from best fit plane RMSi - Residual RMS from best fit sphere after best fit plane has been subtracted RMSa - Residual RMS from best fit AAS after subtraction best fit plane and sphere
E	sub-aperture size	The specifications may be applied to a sub-aperture of the actual optics clear aperture. This would be read as "applies to any sub-aperture of this size which can fit in the clear aperture". The units are millimeters.

These parameters and the associated data/surfaces are summarized in Table 1. The order of operations is important to insure correct values of the tolerances, especially when using non-circular clear apertures.

"Surface"		PV	Surface RMS	Residual RMS
Measured interferometric cavity	W	-	-	-
Best fit plane to W	P	-	-	RMSt (D)
Total Interferometric Error	TIE = W - P	-	RMSt (D)	-
Best fit sphere	S	Sagitta (A) (Flats only)	-	RMSi (D)
Irregularity Error Function	IRR = TIE - S	Irregularity (B)	RMSi (D)	-
Approximating Aspheric Surface (Rotationally symmetric component of IRR)	AAS	Rotationally Symmetric Irregularity (C)	-	RMSa (D)

Table 1. Tolerance quantities and related surfaces used in the ISO 10110-5.

All tolerance quantities are optional. If either A or B are omitted, they are replace by a "-". If C is omitted, so is the "/". If all of A, B and C are omitted, the "(-)" is omitted. If D is omitted, so is "RMSx <". If there is no sub-aperture specified then (all ϕ E) is omitted. In general, something will be left out. Specifying all of A, B, C and D risks over specifying the surface. Reasons for omitting certain tolerances include:

A	RoC for a spherical surface toleranced in the R specification (see ISO 10110-10)
B	Irregularity is either unimportant, or being controlled by an RMS tolerance
C	No special consideration for spherical aberration or zonal errors is needed
D	This optic is not to be tested using digital interferometry, or the irregularity of this optic is large enough that noise in a digital interferometer will not affect PV (irregularity)
(all ϕ E)	The specification is to be applied to the entire clear aperture.

Examples

Some examples of indicating a specification are shown here. These are excerpted from reference 2.

3 / -	The surface is not specified for figure.
3 / - (2.0)	There is no sagitta error tolerance. If the surface is spherical, then the RoC may be toleranced as per ISO 10110-10. If the optic is flat, then there is no tolerance controlling the power in the surface. The surface is specified for 2.0 fringes of irregularity @ 546.1 nm (= 0.86 λ @ 633 nm). Neither the RSI, nor any RMS error have been toleranced. This specification can be measured with either test plates or an interferometer.
3 / - (0.2) (all ϕ 10)	This example is similar to the previous one, except that the irregularity is now specified as 0.2 fringes. Visual evaluation will be difficult for the 0.2 specification and therefore digital analysis is recommended. The specification is also specified for any 10mm sub-aperture within the clear aperture. This type of specification can be useful if the ray bundle for any field point has a footprint considerably smaller than the surface clear aperture. An example of this is a scan lens.
3 / 3.0 (-) RMSi < 0.05	The sagitta error is specified as 3.0 fringes @ 546.1 nm (= 1.29 λ @ 633 nm) and the RMS irregularity (relative to a sphere fit) is 0.05 fringes @ 546.1 nm (= 0.022 λ @ 633 nm). The RMS specification requires the use of a digital interferometer. However, the sag error specification cannot be measured with the interferometer directly. The optician and QC metrologist can either use test plates for this measurement, or measure the RoC of the surface (using an interferometer or spherometer) and convert the result to a sag error using the equation, $$A = \frac{2\,\Delta R}{\lambda}\left\{1 - \sqrt{1 - \left(\frac{D}{2R}\right)^2}\right\} \approx \frac{\Delta R\, D^2}{4\,\lambda\, R^2} \qquad (1)$$ where the approximation is valid if the optic R/number is large (D/R is small). N is the sagitta error, R is the nominal radius of curvature, ΔR is the radius of curvature deviation from nominal, D is the clear aperture, and λ is the wavelength of the specification (546.1 nm by default).
3 / 5.0 (1.0 / 0.5)	The optic is specified with a sagitta error of 5.0 fringes @ 546.1 nm (= 2.16 λ @ 633 nm). The irregularity is 1.0 fringes @ 546.1 nm (= 0.43 λ @ 633 nm) and an aspheric component of 0.5 fringes @ 546.1 nm (= 0.22 λ @ 633 nm). This optic can be evaluated with either test plates or an interferometer. The sag error specification cannot be measured with the interferometer directly. The optician and QC metrologist can either use test plates for the sag measurement, or measure the RoC of the surface (using an interferometer or spherometer) and convert the result to a sag error using equation 1.
3 / - RMSt < 0.05	This optic is specified with a total RMS error (relative to a plane fit) of 0.05 fringes @ 546.1 nm (= 0.02 λ @ 633 nm). This requires use of a digital interferometer. The use of RMSt instead of RMSi means that this optic is a flat and the sag error is included with the irregularity in this RMS specification.

5. Ambiguities and Peculiarities

Though ISO 10110-5 goes a long way to standardizing the specification of surface form error, there are a number of issues relevant to digital interferometry which are dealt with vaguely, or simply not addressed. Most of these have direct impact on measurement accuracy and machine to machine correlation.

Visual vs. digital

When faced with the option of using visual or digital analysis, the standard only states that RMS specifications require digital interferometry. Reliable measurements for tight tolerances also require digital interferometry but no guidelines are given in the standard or its annexes regarding when visual analysis should be rejected in favor of digital interferometry. It is appropriate to specify digital analysis in a note on the drawing.

Reference surface quality

No mention of measurement accuracy is made in the standard or its appendices. While this might not normally be expected in a standard, ISO 10110-5 is written in the language of a specific test procedure - interferometry - and yet no mention is made of the accuracy of the reference surface used. For example, the standard might require that the surface form errors in the reference surface be no more than 10% of those called out in the drawing. This type of requirement is more difficult to achieve as the tolerances on the part become tighter, and yet this fundamental source of error must be accounted for when specifying and qualifying high quality optics. No mention is made in the standard of absolute testing techniques.

Spatial frequency bandwidth

Spatial sampling is a measurement parameter completely omitted in ISO 10110-5. Static fringe analysis and PMI will yield different results. This tends to affect the irregularity tolerance more than the RSI and affects the sagitta and RMS tolerances least. Once again, if this if of concern to the designer, spatial sampling should be specified in a note on the drawing. (PMI may sample higher densities than wanted. It may be more meaningful to ignore spatial frequencies, for example, but the standard does not provide for this.)

Calculation of the AAS

The exact procedure for deriving the AAS is not specified in the standard. The method suggested in the annex is informative only, and states a minimum number of Zernike polynomials to use.

1 Fit a 36 term Zernike polynomial to the Irregularity function (ordering of polynomials specified in ISO 10110-5, section A1.3)
2 Generate a surface map using coefficients 8, 15, 24 and 35 from the Zernike fit (all others set to zero)
3 Mask the generated surface to match the region of real data (necessary only for non-circular apertures)
4 Find the Peak-to-Valley of this generated surface

This implies that equivalent means are acceptable. One proposed alternative uses a series of circular zones.[3] The zonal calculation replaces each single pixel wide zone with the average surface height of the irregularity function in that zone. This method provides uniform spatial resolution of the AAS in the radial direction. It is also easily implemented with phase measuring interferometry, but not with static fringe analysis. Fig. 1 shows a surface irregularity function, the AAS derived from 36 Zernikes and the AAS from 16 and 82 zone calculations. A comparison of these two methods is shown in Fig. 2 where the zonal RSI for a variety of surfaces is shown as a function of the number of zones. The labeling in the legend indicates the dominant error in the part. In each case, the zonal RSI value has been normalized by the RSI obtained by the 36 term Zernike fit. One would expect the RSI value to approach a limit and the measurements bear this out. Somewhere in the neighborhood of twenty zones is sufficient for the sample cases shown. Note that some surface functions yield a zonal RSI which is 20-30% greater than that of the Zernike fit. These differences warrant further evaluation to determine whether or not it may be necessary to specify the particular algorithm used for the AAS and number of Zernikes or zones, in a note on the drawing.

Aspheres

ISO 10110-5 is primarily directed at plano and spherical surfaces. It is also extended to aspheres through two rather terse footnotes. The procedures deduced from these footnotes are as follows:

Sagitta error

1. Measure an aspheric part.
2. From the data, subtract the best fit plane.
3. Then subtract the aspheric surface specified for the optic to get the Total Interferometric Error function.
4. From the TIE, find the best fit sphere.
5. Find the PV of the best fit sphere, this is the sagitta error.

Irregularity

1. Measure an aspheric part.
2. From the data, subtract the best fit plane.
3. Then subtract the aspheric surface specified for the optic to get the Total Interferometric Error function.
4. From the TIE, subtract the best fit sphere. to get the Irregularity function.
5. Find the PV of the Irregularity function, this is the Irregularity.

These footnotes make no reference to visual vs. digital evaluation, though step 3 in each case precludes the use of visual analysis and implies that aspherics must be measured using some type of digital interferometry. There is no reference made to accuracy concerns when measuring even mildly aspheric surfaces with interferometers. Specification of aspheric surfaces is also dealt with in section 12 of the standard.

Non-circular apertures

Annex 1 of the standard recommends using Zernike polynomials for several of the calculations. This is fine for optics which are completely circular and which have surface form errors which resemble the "primary" aberrations (spherical, coma and astigmatism) If these conditions are not met, then one must be a bit more careful in calculating and interpreting the results.

A non-circular aperture means an aperture which is elliptical, or polygonal, or has an obscuration, or some irregular shape. In these cases, the Zernike polynomials are no longer orthogonal and will not behave as expected. The order in which reference surfaces are fit and subtracted becomes important; first plane, then sphere, and finally AAS.

The calculation of Sagitta, Irregularity, RMSx, and RSI must be done on a point by point basis for non-circular apertures, as mentioned in section 3.6 of the standard. That is, the RMSx should be calculated as residual RMS (see the Glossary for details). Also the Irregularity and RSI must be calculated from the actual data points in the part clear aperture, relative to the best fit sphere and AAS, respectively. Using the Zernike polynomials to calculate sagitta error for non-circular apertures is still valid if the fit is limited to the first four polynomials.

Other Zernike problems

The Zernikes polynomials may not be able to adequately represent the surface form error of the optic under test. This can happen, for example, when the optic has a pronounced intermediate zone or a "hole" in the middle , or some other type of localized surface error.. It can be useful to examine the residual RMS values for the Zernike polynomial fit in order to judge how well the surface is represented by the polynomials.

Visual vs. digital (again)

Despite its efforts, the standard has still left some differences between visual and digital analysis. For example, the irregularity is the PV of the surface relative to the best fit sphere. For visual analysis, the irregularity is the astigmatism (cylinder) component of the irregularity function. This will cause different results for parts with zonal or local surface errors. There is potentially a similar discrepancy for RSI. If the visual analysis does not properly estimate the AAS, then the visual RSI may be closer to the PV of the irregularity function relative to a sphere fit.

Sub-aperture specification

The sub aperture specification is also somewhat ambiguous. There is no statement regarding the number of sub-apertures to be tested or their distribution within the clear aperture. This is a good example of where a note on the drawing will be useful.

What ISO 10110-5 does not specify

There are a number of parameters which are used to specify optics which are not part of ISO 10110-5. If these are used, they must be called out separately in a note on the drawing. These include: surface slope, PSF, MTF, encircled energy, and Strehl ratio. Of these evaluation criteria, surface slope is most often used for specification of surfaces. This is not currently standardized and is a frequent source of confusion when it appears on an optical drawing.

ISO 10110-5 does not specify surface form error tolerances at non-normal incidence angles. For example, a fold mirror is specified as if it were used at normal incidence, not at some other angle. It is the responsibility of the designer to relate the requirements to the normal incidence specification, or to specify a non-normal test and specification in a note on the drawing. This would also be required for conic surfaces tested at their conjugates.

It should be mentioned again that ISO 10110-5 applies only to surfaces, and is not applicable to transmitted wavefront testing as for lenses, window and optical assemblies.

6. Summary

ISO 10110-5 represents a good first step to providing standards in a relatively non-standardized industry. There are several aspects of the standard which are in need of further specification for digital interferometry. These include:

- spatial sampling bandwidth
- calculation algorithms for the AAS
- specifics of sub aperture testing
- accuracy of reference surfaces
- a standard for slope specification

It presents a challenge to users and manufacturers of digital interferometers to propose test and debate further standardization to fill the voids in the ISO standard.

Acknowledgments

Thanks to Dr. John Rogers of Wild-Leica and Gordon Kane of Computer Optics, Inc. for their discussions and suggestions. Thanks also to Peter Pfluke for software and measurements.

References

1. ISO/DIS 10110-5, revision April 18, 1992
2. Kane, G. and L. A. Selberg, "Surface Form Tolerances" in OSA Handbook of Optical Standards, R. E. Parks, and R. Kimmel, ed., Optical Society of America, 1992 (in press)
3. Rogers, J. R., private communication, 1992

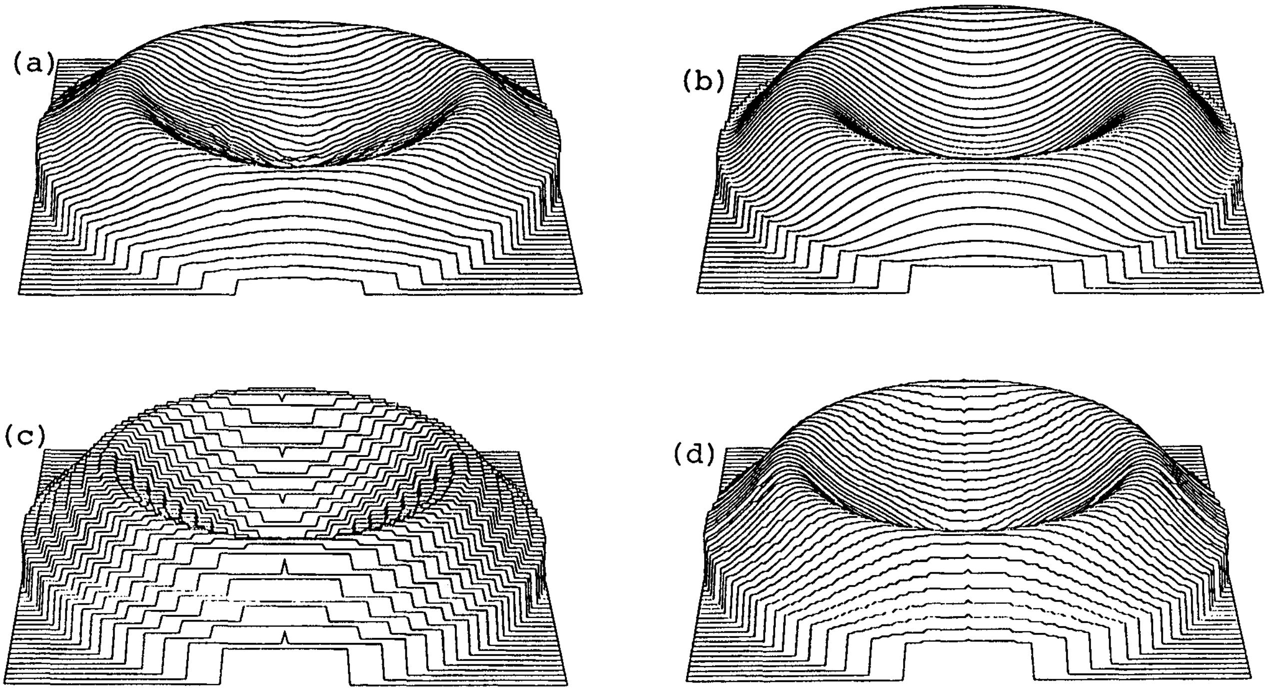

Figure 1. (a)The irregularity function for a surface with a "hole and roll". The AAS is shown as calculated from (b) a 36 term Zernike fit, (c) a sixteen zone calculation, and (d) an 82 zone calculation.

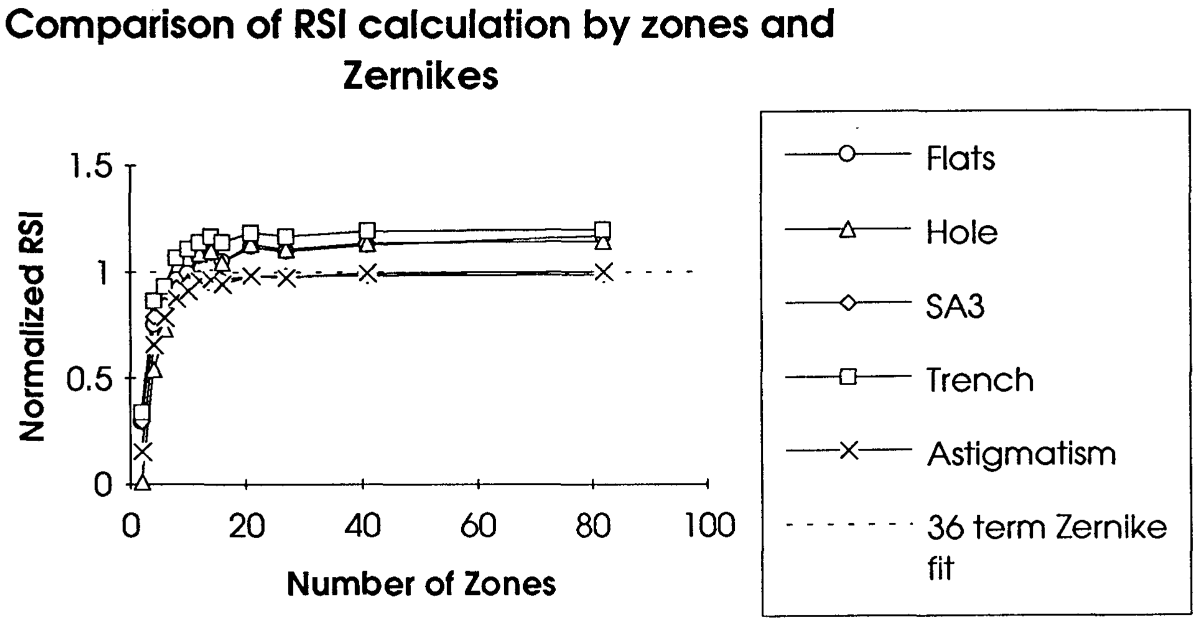

Figure 2. RSI for a variety of different surfaces calculated by a zonal technique, shown as a function of the number of zones. In each case, the zonal RSI is normalized by the RSI from the 36 term Zernike calculation. The surfaces are labeled to correspond with the dominant surface form error.