Roughness Measurement Performance Obtained with Optical Interferometry and Stylus Method

Hyug-Gyo Rhee*, Yun Woo Lee, and In Won Lee

Korea Research Institute of Standards and Science, Space Optics Research Center, 1 Doryong-dong, Yuseong-gu, Daejeon, 305-340, KOREA

Theodore V. Vorburger

National Institute of Standards and Technology, Precision Engineering Division, Surface and Microform Metrology Group, NIST 100 Bureau Dr. Stop 8212, Gaithersburg, MD20899, USA

White-light scanning interferometry (WLI) and phase shifting interferometry (PSI) are increasingly used for surface topography measurements, particularly for areal measurements. In this paper, we compare surface profiling results obtained from above two optical methods with those obtained from stylus instruments. For moderately rough surfaces ($Ra \approx 500$ nm), roughness measurements obtained with WLI and the stylus method seem to provide close agreement on the same roughness samples. For surface roughness measurements in the 50 nm to 300 nm range of Ra, discrepancies between WLI and the stylus method are observed. In some cases the discrepancy is as large as 109% of the value obtained with the stylus method. By contrast, the PSI results are in good agreement with those of the stylus technique.

OCIS codes : 120.3180, 120.2830, 120.6650, 120.6660, 180.3170

I. INTRODUCTION

Recently, the International Organization for Standardization (ISO) formed a working group for standardization of areal (3D) surface texture measurement methods as well as review existing standards on traditional profiling (2D) methods. A number of draft standards are being developed. In one of these, surface texture measurement methods are classified into three types, line profiling, areal topography, and area-integrating [1,2]. A line profiling method uses a high resolution probe to sense the peaks and valleys of the surface and produce a quantitative profile z(x) of the surface topography. Areal topography methods extend the line profiling method into three dimensions usually by some quantitative topographic imaging process. Area-integrating methods probe an area of the surface all at once to produce a quantity that represents some statistical average of the surface peaks and valleys.

In this paper we will focus on areal topography methods and in particular on the widely used phase shifting interferometry [3] (PSI) and white-light scanning interferometry [4-6] (WLI). WLI seems practically versatile because PSI is limited to measurement of smooth surfaces, whereas

the vertical dynamic range of WLI extends from the nanometer level (noise level) to a large range. However, WLI has some problems with step heights [7] and rough surfaces [8] whose peak-valley values are less than the coherence length of the light source. To observe the problems, we compared roughness measurements obtained from two optical techniques with those obtained from stylus instruments. The stylus instrument serves as a standard technique for these comparisons. For moderately rough surfaces ($Ra \geq 500$ nm) and fine surfaces ($Ra \leq 30$ nm), roughness measurements obtained with WLI and the stylus method seem to provide close agreement on the same roughness samples. For surface roughness measurements in the 50 nm to 300 nm range of Ra, discrepancies between WLI and the stylus method are observed. In some cases the discrepancy is as large as 109% of the value obtained with the stylus method. By contrast, the PSI results are in good agreement with those of the stylus technique.

II. SAMPLES MEASURED

For the roughness measurements, we used the roughness average parameters, Ra and Sa [9], for comparison

of the measured topographic results. For a digitized profile Ra is defined as

$$Ra = (1/N)\sum_{i=1}^{N}|Z_i| \ , \qquad (1)$$

where Z_i is the series of N profile data points. Analogously, for a surface topography image, Sa is defined as

$$Sa = \left[(1/NM)\sum_{i=1}^{N}\sum_{j=1}^{M}|Z_{ij}|\right] \ , \qquad (2)$$

where Z_{ij} is the array of $N \times M$ points in a topographic image. When the measured surfaces are highly uniaxial (i.e., the machining marks are oriented along one direction), Sa results calculated from topographic images should be very close to Ra results calculated from profiles.

We tested standard periodic gratings with Ra ranging from about 60 nm to 500 nm and random roughness standards with Ra values ranging from the sub-nanometer level (i.e., a smooth mirror) to about 130 nm. Except for the smooth mirror, the surfaces are highly uniaxial.

These standards are listed in Table 1.

The mirror has a smooth isotropic surface with an Sa value approximately equal to the resolution limit of PSI. The mirror is especially useful for estimating the resolution limit for the other optical techniques. The four random samples, Rubert 501-504 [10], are electroformed Ni replicas of steel surface roughness samples fabricated by Song [11] using a hand-lapping process. In addition to being highly uniaxial, the "random" roughness pattern is repeated over distances equal to the recommended profile evaluation length so that the roughness value is highly uniform when measured in widely different places.

The four periodic surfaces came from three different sources. Two of these (Rubert 528, 529) are electroformed Ni replicas of sinusoidal profile roughness samples machined by numerical controlled diamond turning. NIST Standard Reference Material (SRM) 2071 [12] has an electroless Ni surface, diamond turned to produce a sinusoidal profile. The sinusoidal holographic grating has a Au coating to provide a surface with a uniform reflectivity.

TABLE 1. Samples used for the roughness measurements. The uncertainties in column threeare for a coverage factor k = 2.

Sample Identification	Nominal Ra Value (nm)	Measured RaValue using Stylus (nm)	Measured $R\Delta q$ value using stylus[d] (mrad)	Other Specifications
Smooth Mirror[a]		⟨0.16 (measured with PSI)		
Rubert[c] 501[a]	15	14.0 ± 1.8	37.4	
Rubert 502[a]	35	34.5 ± 3.4	38.3	
Rubert 503[a]	90	88.4 ± 5.8	83.9	
Rubert 504[a]	130	129.7 ± 8.6	108.5	
Au-coated Holographic grating[b]	60	58.9 ± 6.8		Nominal 6.6 μm period
Rubert 529[b]	105	103 ± 10	78.2	Nominal 10 μm period
NIST SRM 2071[b]	300	321.7 ± 8.7	42.8	Nominal 100 μm period
Rubert 528[b]	502	506 ± 29	78.4	Nominal 50 μm period

[a] Random profile samples.

[b] Sinusoidal gratings.

[c] Certain commercial materials are identified in this paper to specify adequately an experimental procedure. Such identification does not imply recommendation or endorsement by KRISS and NIST, nor does it imply that materials are necessarily the best available for the purpose.

[d] Slope measurements of surface profiles depend sensitively on the sampling interval used. Therefore absolute uncertainties are difficult to assess. Because all of the $R\Delta q$ value were calculated with the same sampling interval, the relative uncertaintiesdue to position variation are approximately less than 5% (k=1).

III. MEASUREMENT CONDITIONS

As stated before, the stylus instrument served as a standard measurement for these comparisons. Therefore, except for the smooth mirror, all surfaces were measured with a stylus instrument, and the Sa results obtained with the two types of optical microscopes were compared with the Ra results obtained with the stylus. Stylus profiling of the smooth mirror might scratch the mirror and damage its primary function, so PSI serves as the reference technique for the mirror because its noise resolution limit is smaller than that of WLI. For all samples, the surfaces were measured at several different positions on the surface to assess the stability of the roughness values as a function of position. Results for roughness average Ra were calculated from the surface profiles and are shown in Table 1. The expanded (k =2) uncertainties[13] of these Ra values are comprised of several components, a Type A statistical variation [13] and Type B components due to z-scale calibration, z-scale resolution, uncertainty in the long and short-wavelength cutoffs, and uncertainty in the stylus radius. The uncertainty of the Sa value obtained by PSI for the smooth mirror is such that we quote an upper limit for Sa in Table 1. Because the surfaces are fabricated to function as roughness samples, they have very low waviness and flatness deviations. Therefore, Ra values calculated from unfiltered profiles are not significantly different from the values shown in Table 1.

Several stylus instruments were used. For the random surfaces and the Rubert 528 grating, a band-pass Gaussian filter [14] (cutoffs: 250 μm and 2.5 μm) was used. For the Rubert 529 grating, the long-wavelength cutoff was 800 μm and the short-wavelength cutoff was 2.5 μm. SRM 2071 was calibrated with a different stylus instrument having a 2RC [14] filter with a long-wavelength cutoff of 800 μm. Because the sinusoidal profile has a spatial wavelength of 101.6 μm, the Ra value that would be obtained with a Gaussian filter with 250 μm cutoff is estimated to be within 1% of the value shown in Table 1. SRM 2071 is used periodically as a roughness check standard in the NIST and has been

shown to have a stable roughness value. The nominal stylus radius was 2 μm. The Au-coated holographic grating has been tested with several instruments under different conditions. The nominal stylus radius was 2 μm or smaller. The Ra value of 58.9 nm shown in Table 1 was measured with a long-wavelength cutoff of 250 μm.

We also performed PSI microscopy using a commercial instrument with the Mirau type configuration. For all surfaces except SRM 2071, a 50× objective was used having a numerical aperture (NA) of approximately 0.55. The field of view on the surface was approximately 120 μm × 90 μm. For SRM 2071, a 20× objective was used, with a field of view of approximately 300 μm × 230 μm. The instrument calibration was checked with a calibrated step standard having a 90.65 nm height. For PSI and for the other microscopes as well, Sa values were calculated from the topographic images without digital long-wavelength filtering, partly because the topographic images had small fields of view. The field of view along the roughness direction determines the evaluation length and essentially provides the long-wavelength cutoff limit. For all surfaces except the 503 and 504, average values were calculated from Sa results measured in at least three positions. The 503 and 504 standards are difficult to measure with PSI because of their relatively high roughness and short spatial wavelengths. Measurements from only two positions were used for the 504 and only one was used for the 503. Other measurement attempts on these surfaces produced obvious unwrapping errors [3], which were observed in the topographic images as z-discontinuities producing surface patches with different apparent heights.

Moreover, we tested five different WLI microscopes as shown in Table 2. One of them (Instrument A) is the same instrument as that used for PSI. All WLI microscopes were used in the Mirau type configuration. For instruments A and B, 50× objectives were used to measure all surfaces except SRM 2071 for which 20× objectives were used. The fields of view of the topographic images were the same as for PSI. All nine

TABLE 2. Specifications of five optical instruments.

	Instrument 1	Instrument 2	Instrument 3	Instrument 4	Instrument 5
Detector type	736×480 CCD	736×480 CCD	640×480 CCD	640×480 CCD	512×512 CCD
Nominal sampling space	0.413 μm with 20× objective 0.165 μm with 50× objective	0.413 μm with 20× objective 0.165 μm with 50× objective	0.45 μm with 20× objective 0.18 μm with 50× objective	0.55 μm with 20× objective 0.22 μm with 50× objective	0.55 μm with 20× objective 0.22 μm with 50× objective
Vertical resolution	3 Å for PSI 3 nm for WLI with single measurement	3 Å for PSI 3 nm for WLI with single measurement	less than 1 nm for WLI with 5 times average	less than 1 nm for WLI with 5 times average	4 nm for WLI with single measurement

samples were measured with instrument A and seven of them, Rubert 501-504, 528, and 529 and SRM 2071 were measured with instrument B. For instruments C, D, and E, four samples were measured, Rubert 502, 504, 528, and 529.

IV. EXPERIMENTAL RESULTS

Figs. 1, 2, and 3 show the differences in measured roughness average (Sa -Ra) values between the optical techniques and the stylus technique.

Fig. 1 shows generally good agreement between the PSI results and the stylus results for the range of smooth surfaces where PSI is expected to be accurate. These surfaces include the random roughness surfaces with Ra ranging up to 130 nm, and the sinusoidal surfaces with Ra ranging up to 500 nm. The result for the smooth mirror was not plotted because the stylus was not used to measure it, as previously discussed. The Sa - Ra differences are less than or equal to 6 nm, except for the Rubert 504 and 503 samples with nominal Ra of 130 nm and 88 nm, which have surface slopes that may be high enough to cause inaccuracy in the measured roughness topography by PSI because of phase-wrapping issues. High slopes lead to large height differences between pixels, which in turn can cause phase wrapping ambiguities and hence height measurement errors in the surface profiles.

As stated before, only one topographic image was included in the results for the 503 sample because phase unwrapping errors were observed in the other images. For the rest of the data, the variation in the measured data with position was such that typical standard deviations of the mean were 2.7 nm.

Fig. 2 shows larger discrepancies between WLI and the stylus technique than Fig. 1 for the PSI. The discrepancies exist for both random and periodic surfaces

and seem to peak in the Ra range between about 100 nm and 200 nm. The largest of the observed discrepancies is about 80 nm or about 109% of the Ra value for the 60 nm Ra standard as obtained with the stylus technique. Typical standard deviations of the mean of the measured Ra values are about 6 nm here. Results for five different instruments are shown in Fig. 2, two different models from the same manufacturer (Instruments A and B), the other two models from a second manufacturer (Instruments C and D) and a fifth instrument from a third manufacturer. Instrument E uses a totally different WLI algorithm from others. The relation between the discrepancy and the Ra value, therefore, is not related to a specific WLI instrument. In addition, the shape or the randomness of the sample is not an important factor because the differences are observed for both random and periodic surfaces. The differences do not seem to be present if the surface is sufficiently rough or smooth, but seem to be caused by issues in the instrument optics when the surfaces are in a middle range of roughness such that they have a somewhat glossy appearance. Note that for the smooth mirror, the Ra value obtained with WLI is 3 nm, consistent with the PSI value of 0.16 nm, given the resolution limit of WLI.

The cause of these effects is not yet known to us. We are convinced that the differences illustrated in Fig. 2 are due to the WLI, because the PSI readings seem to have no special relation to the Ra value in Fig. 1.

An important issue is any sensitivity of the Ra results to the field of view of the microscopes. For the periodic samples, this is not expected to be significant at all and only very small differences were obtained when the objective was changed from 50× to 20×. For the random surfaces, both PSI and WLI measurements with the 50× objective yielded topographic images with fewer data dropouts (invalid points) than with the 20× objective. That is because the 50× objective has a higher numerical aperture and can measure local areas with higher surface slopes than the 20× objective. Also,
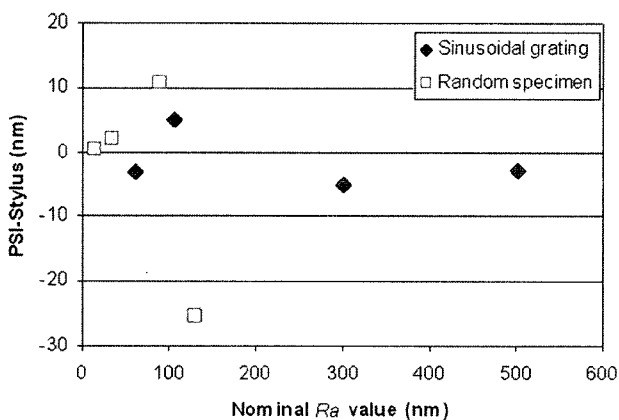

FIG. 1. Differences of measure *Ra* values obtained with the PSI and the stylus technique as a function of the nominal *Ra* value for four random roughness samples and four sinusoidal gratings.
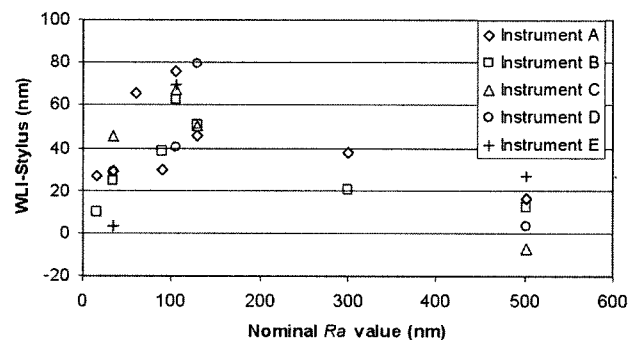

FIG. 2. Differences of measure *Ra* values obtained with WLI and the stylus technique as a function of the nominal *Ra* value. Results are shown for five WLI microscopes.

because the point spacing is closer, phase unwrapping errors are expected to be smaller for the 50× objective. When we switched from the 50× to the 20× objective on the random samples, in most cases, the measured roughness got slightly smaller. This is surprising because increasing the field of view generally increases the measured roughness amplitude for a random roughness surface. The fact that this did not happen with increasing field of view suggests that surface structures observed at 50× were not observed at 20× because of a loss of spatial resolution.

Fig. 3 shows a portion of four profiles measured for the 105 nm Ra Rubert 529 standard with each of the three profiling techniques. Although the surface is clearly periodic, the sinusoidal profile at this fine roughness level is imperfect. Despite the distortion, we can observe clear differences between the profiles measured for this surface. The stylus and PSI profiles have about the same amplitude, but the spatial frequency content of the stylus profile seems to be higher whereas the PSI profile seems to be smoother. The WLI profile has higher amplitude, is clearly distorted, and shows noise-like spikes on the grating peaks.

Fig. 4 shows profiles measured with the three techniques for the Rubert 528 sinusoidal grating, which has 500 nm Ra and 50 μm spatial wavelength. The discrepancies between the profiles are now significantly smaller relative to the scale of the surface profile structure than they were for the Rubert 529 grating.

V. DIFFRACTION MODEL

A model to describe the discrepancy of WLI was already suggested by the authors [8]. In the model, we assumed a normal incidence plane wave on a sinusoidal grating surface. We also assumed that the Fresnel app-
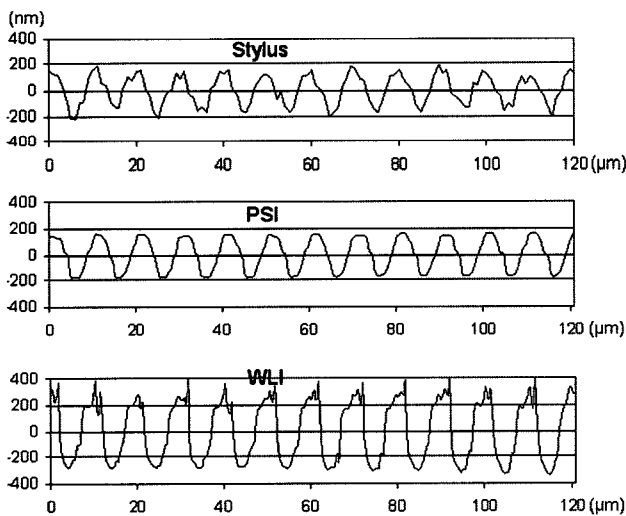
roximation was not valid in our case because our interest was focused on a sub-micrometer level surface variation. For simplicity we illustrated the model in one dimension only, which means that the sample varies in one direction. This aspect of the model is consistent with the samples we used as shown in Fig. 5.

The objective lens collects not only the reflected/diffracted beam from the ideal imaging point but also the neighboring light from the vicinity within a 1.22λ/NA diameter. In addition, the CCD camera has finite pixel size, approximately 9.8 μm × 8.4 μm, so that the measured intensity is obtained from the sum of incidence light in a pixel. This neighboring light might influence the errors in a WLI measurement. Particularly when the surface variation is less than the depth of focus of the objective lens and less than the coherence length, the interference among these beams will be more apparent.

Here in this paper, we slightly changed the simulation conditions. We used the smallest wavelength of the light $\lambda_1 = 350$ nm, the biggest wavelength $\lambda_2 = 650$

FIG. 4. Profiles of the 500 nm Ra sinusoidal grating obtained with three techniques.
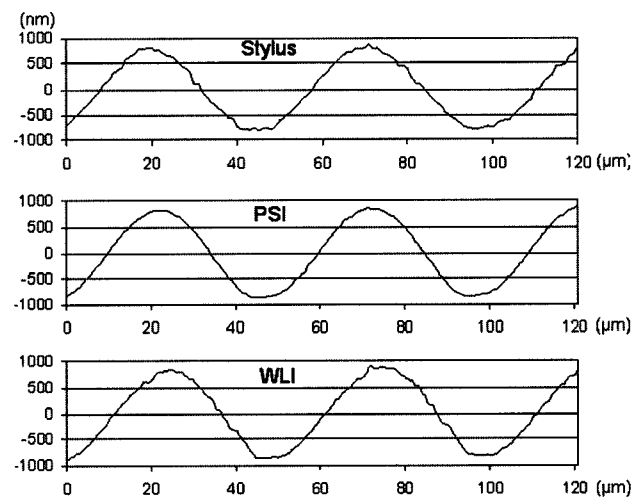
FIG. 3. Profiles of the 105 nm Ra sinusoidal grating obtained with three techniques.
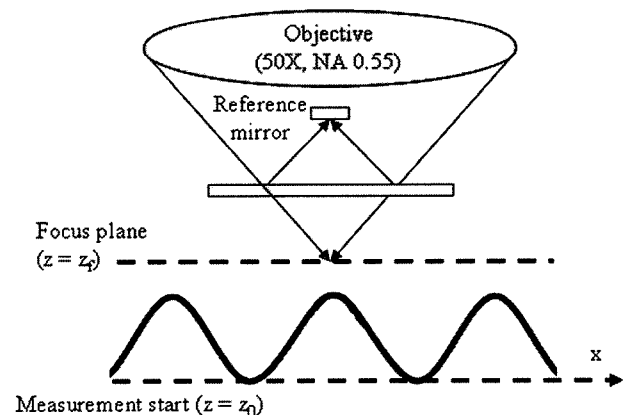
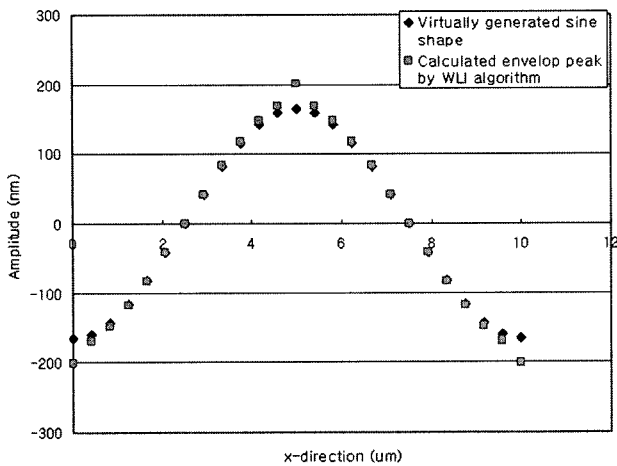FIG. 5. Configuration of a modeled interferometer system.

FIG. 6. Simulated WLI readings with a virtual sine shape. The virtual sample was generated with 330 nm peak-valley value and 10 μm surface spatial wavelength.

nm, the center wavelength $\lambda_0 = 500$ nm, and the vertical sampling distance $\Delta z = 80$ nm. For numerical calculation we took 1000 data points in the horizontal direction. The test sample was a virtual sinusoidal surface profile with 330 nm peak-valley value, 10 μm surface spatial wavelength and about 105.04 nm in Ra value as shown in Fig. 6.

It is the same as the Rubert 529 grating that yields the largest discrepancy among the tested sinusoidal gratings. After simulation, 401.23 nm peak-valley value was calculated with the WLI algorithm. We can also see the noises at the top and the bottom and relatively small distortion elsewhere. At the side of the sine shape, the nearby vicinity light exerts an influence on the interference pattern too. However, its effect appears symmetrical in the interference pattern so that the envelope peak of the pattern is not affected. One thing to be noted is that the simulation is not a perfect match for the experimental results in the chapter IV. It may be that the one dimensional simulation conditions we assumed are not perfect, or that the error arises from a different source. There are many other possible error sources such as nonlinearity or slope effects. Further, we may need an extension of our model that includes nonlinearity, NA effects, light scattering, and the slope of the surface feature.

V. DISCUSSION AND CONCLUSION

We have tested periodic grating standards and random roughness standards, whose Ra values range from sub-nanometer levels to 500 nm. Within this range, the discrepancy between WLI and the stylus instrument shows a special correlation with the surface roughness parameter Ra. This discrepancy is especially prominent for Ra values between 100 nm and 200 nm, a transition

range of roughness between smooth, specular surfaces and rough, diffuse surfaces. The discrepancy seems to be unrelated to the specific instrument, specific WLI algorithm, specimen shape, and randomness.

In Fig. 2, we have plotted the Sa - Ra values for WLI as a function of the Ra value. However, it is not yet clear what surface parameters correlate most closely with the measured deviations. Roughness amplitude is clearly important, and the Ra value is a good metric of average amplitude for both random and periodic surfaces. Other averaging roughness-amplitude parameters, such as the root mean square roughness (Rq), are also useful measures. There does not seem to be a correlation with root mean square slope $(R\Delta q)$ of the surface. One indication of this is that the 100 nm and 500 nm sinusoidal surfaces have nominally the same slope but exhibit markedly different profile distortion with respect to the stylus profile. Perhaps surface curvature or peak curvature is the combination of amplitude and wavelength properties that most directly affects the deviations shown in Fig. 3.

The deviations in the measured profile and Ra values for WLI may be related to a similar effect observed for step height measurements. WLI has been shown to be susceptible to a skewing effect for step height measurements when the amplitude is less than the coherence length of the light source [7]. The skewing effect leads to spikes in the profile data near step edges. Accurate results for step heights may be obtained by avoiding the edge regions in step-height analysis [8]. Analogously for roughness, regions of high curvature at surface peaks and valleys may produce similar profile distortions, as seen for the WLI profile in Fig. 3. A diffraction model is also presented to describe the phenomenon.

*Corresponding author : hrhee@kriss.re.kr

REFERENCES

[1] T. V. Vorburger, J. A. Dagata, G. Wilkening, and K. Iizuka, *Beam Effects, Surface Topography, and Depth Profiling In Surface Analysis* (Plenum, New York, USA, 1998), pp. 275-354.

[2] T. V. Vorburger and J. Fu, N. Orji, "In rough," *OE Magazine*, March 2002, pp. 31-34, 2002.

[3] J. E. Greivenkamp and H. Bruning, *Optical Shop Testing* (Wiley, New York, USA, 1992), pp. 501-598.

[4] L. Deck and P. de Groot, "High-speed noncontact profiler based on scanning white-light interferometer," *Appl. Opt.*, vol. 33, no. 31, pp. 7334-7338, 1994.

[5] A. Harasaki, J. Schmit, and J. C. Wyant, "Improved vertical-scanning interferometry," *Appl. Opt.*, vol. 39, no. 13, pp 2107-2115, 2000.

[6] M. G. Kang, S. Y. Lee, and S. W. Kim, "Self-compensation of PZT errors in white light scanning interferometry," *J. Opt. Soc. Kor.*, vol. 3, no. 2, pp. 35-40, 1999.

[7] A. Harasaki and J. C. Wyant, "Fringe modulation

skewing effect in white-light vertical scanning interferometry," *Appl. Opt.*, vol. 39, no. 13, pp. 2101-2106, 2000.

[8] H. G. Rhee, T. V. Vorburger, J. W. Lee, and J. Fu, "Discrepancies between roughness measurements obtained with phase shifting interferometer and white-light interferometry," *Appl. Opt.*, vol. 44, no. 28, pp. 5919-5927, 2005.

[9] ASME B46. 1-2002, *Surface Texture [Surface Roughness, Waviness, and Lay]* (An American National Standard, The Am. Soc. Of Mech. Eng., USA, 2003), pp. 6-12.

[10] www.rubert.co.uk, accessed 12 December 2005.

[11] J. F. Song, T. V. Vorburger, and P. Rubert, "Com-

parison between precision roughness master specimens and their electroformed replicas," *Prec. Eng.*, vol. 14, no. 2, pp. 84-90, 1992.

[12] T. V. Vorburger, J. F. Song JF, C. H. W. Giauque, T. B. Renegar, E. P. Whitenton, and M. C. Croarkin, "Stylus-laser surface calibration system," *Prec. Eng.*, vol. 19, no. 3, pp. 157-163, 1996.

[13] ISO, *Guide to the Expression of Uncertainty in Measurement* (International Organization for Standardization, Geneva, Switzerland, 1995), pp. 19-24.

[14] ASME B46. 1-2002, *Surface Texture [Surface Roughness, Waviness, and Lay]* (An American National Standard, The Am. Soc. of Mech. Eng., USA, 2003), pp. 25-29.