

Optical methods of measuring rough surfaces

K. Leonhardt, K.-H. Rippert and H. J. Tiziani

Institute of Applied Optics, University of Stuttgart
Pfaffenwaldring 9, 7000 Stuttgart 80, West Germany

ABSTRACT

The statistical parameters of surfaces to be measured for industrial applications vary over several orders of magnitude. Surfaces with large slopes or edges are particularly difficult to be recorded. Some measuring methods developed in our laboratory are compared and the range of applications are discussed.

For polished and fine ground glass and metal surfaces a heterodyne profilometer with a vertical resolution of 0.5 nm, lateral resolution of 0.6 μm , and large scanning length is discussed. The interferometer can be changed from single- to double-pass operation by rotation of a quarter-wave-plate.

For rougher surfaces a profilometer of the photometric-balance type with resolution $R_q < 4$ nm and dynamic range of 20 μm and an interference microscope with automated fringe evaluation is described. An integral white light roughness sensor covers the roughness range 0.04 μm to 10 μm and measures independently mean roughness and autocorrelation width.

1. INTRODUCTION

There is an increasing need for optical profilometry and surface measuring instruments. For smooth surfaces optical methods have become a powerful tool^{1,2} particularly for high quality optical surfaces. Resolution of 0.1 nm specified as rms-values have been reached. However for surfaces with larger local slopes such as of fine ground surfaces or for surfaces with large roughness values as used in metalworking industries for metallic surfaces, problems of reliance and of large local errors arise. In this paper we describe results for non-smooth surfaces of roughness values in the nm- and μm -range.

2. DEFINITIONS FOR OPTICAL SURFACE AND ROUGHNESS MEASUREMENT

In order to test and to discuss measuring methods and instruments, the test surfaces should be carefully selected. They should be globally flat or geometrically well defined with mean roughness much smaller than the measuring length L and with negligible contributions of spatial frequencies $f < 1/L$. The height profile $h(x)$ can then be described as a stationary random process with mean $\langle h(x) \rangle = 0$ given by the raw data $y(x)$ by subtraction of a best fitted quadratic curve,

$$h(x) = y(x) - a - b \cdot x - c \cdot x^2 \quad (1)$$

(or a straight line for $c=0$). This trend removal is necessary to exclude residual set-up errors.

In this paper we use the rms-value or quadratic mean roughness R_q as an estimate of the standard deviation σ_h of $h(x)$

$$\sigma_h^2 = E \langle h_i^2 \rangle \quad (2)$$

$$R_q^2 = \frac{1}{L} \int_0^L h(x) dx \hat{=} \frac{1}{N} \sum_{i=1}^N h_i^2 = \langle h_i^2 \rangle \quad (3)$$

In (3) ergodicity and negligible sampling errors are assumed. The number of discrete sample values should be large enough to assure stationarity of statistical characteristics.

The correlation function of second order $C(\Delta)$ or the autocorrelation function $r(\Delta)$ provide horizontal and vertical information

$$r(\Delta) = \frac{C(\Delta)}{\sigma_h^2} = \frac{1}{\sigma_h^2} E \langle h(x) h(x+\Delta) \rangle \quad (4)$$

where Δ is in the autocorrelation lag along the x -axis^{3,4}. The autocorrelation function is related to the spectral density function $S(f)$ by Fourier transform:

* We denote the mathematical expectation value by $E \langle \dots \rangle$ and the corresponding measured sample estimate with $\langle \dots \rangle$.

$$S(f) = 4 \sigma_h^2 \int_0^{\infty} r(\Delta) \exp\{-i2\pi\Delta f\} d\Delta = 2\sigma_h^2 \int_0^{\infty} r(\Delta) \cos\{2\pi\Delta f\} d\Delta \quad (5)$$

where f is the spatial frequency and $S(f)$ is the one-sided spectral density or power spectrum. As a horizontal descriptor we use the correlation length w_h as width of the autocorrelation function, within which the surface heights $h(x)$ are correlated. For surfaces with negligible contribution of low frequency-components due to the machining process (non-fractal) the objections⁵ against autocorrelation width do not apply.

3.0 Optical microprofilometry

The surface is scanned relatively to the detection system in the same manner, as it is done by mechanical profilometry. However, the diamond stylus is replaced by a focus of a gaussian beam with the width of 0.3 μm to 5 μm , depending on the numerical aperture of the profiling lens.

3.1 Heterodyne profilometry The principle of this profilometer is based on heterodyne interferometry. Fig.1 shows the arrangement of our heterodyne profilometer^{10,11}. A He-Ne laser with Zeeman splitting of the 633 nm line is used as source. Two orthogonally polarized modes with a beat frequency of 2 MHz are focused onto the surface by the scanning lens O3 and held in focus by autofocus control (AFC). The interferometer is constituted by the polarizing beam-splitter BS, separating object and reference arm.

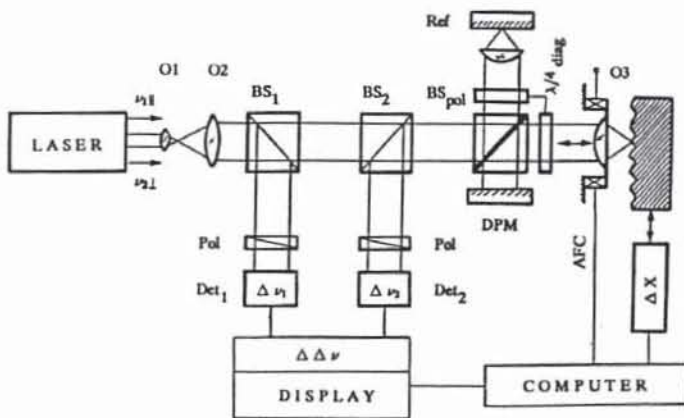


Fig 1: arrangement of the heterodyne profilometer

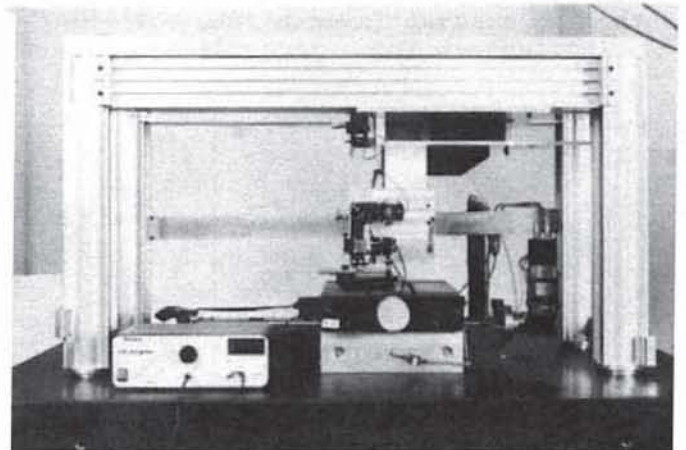


Fig 2: heterodyne profilometer

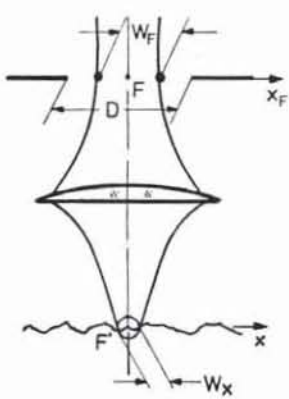


Fig.3 Aperture of the scanning lens

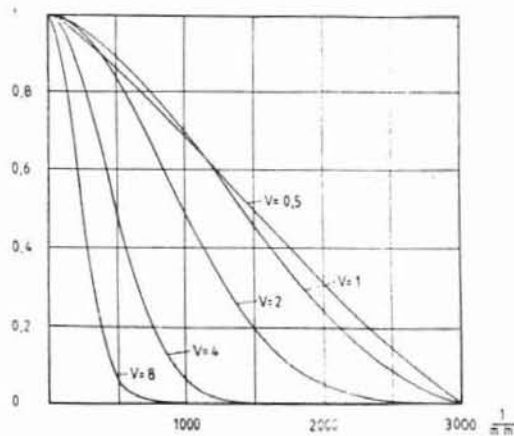


Fig.4 Height transfer function of the heterodyne profilometer

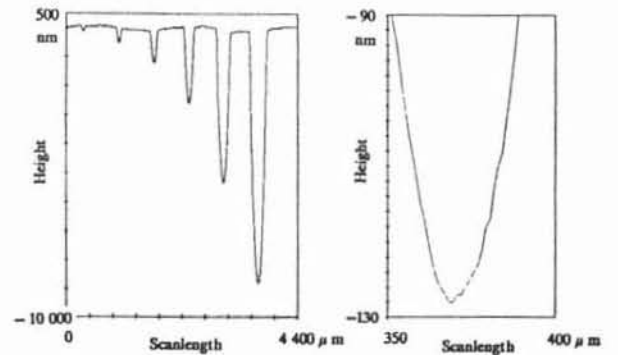


Fig.5a and b: PTB-calibration standard

Microscope lenses with very high aperture $NA=0.95$ can be used. A double-pass system, described in detail in 3.3 consisting of a quarter-wave plate in each arm and a double-pass mirror DPM serves for compensation of local profile slopes. In contrast to^{6,7}, the profilometer is not a differential type but measures profil height directly with reference to an air-bearing slideway of 300 mm length (Fig.2). Thus shape measurements (Fig.5) in the 10 nm to 40 μm range can be obtained directly on smooth surfaces, and roughness measurements in the nm- or sub-nm-range can be obtained after digital high-pass-filtering of the raw-profile

with cutoff-wavelength of typically $100\ \mu\text{m} - 500\ \mu\text{m}$. A very high degree of flexibility is guaranteed from this concept.

The incremental resolution step of our digital electronics is $0.5\ \text{nm}$. Lateral resolution is given by the aperture, $\text{NA}=0.95$, and the aperture ratio V (Fig.3), $V = D/w$, with aperture width D and width of the gaussian beam waist w_F in the focal plane of the profiling lens. The beam waist w_F can be adjusted by the focal length ratio of the beam expander O1, O2. In Fig.4 the profile height transfer function with V as a parameter is depicted as a function of spatial frequency f ($\lambda = 633\text{nm}$). Very high lateral resolution up to the frequency limit of $f = 3000\ \text{mm}^{-1}$ can be realized. Fig.5 shows the profile of a PTB- calibration standard used to calibrate mechanical stylus instruments. Six grooves, precisely lapped into a plane glass plate, are profiled

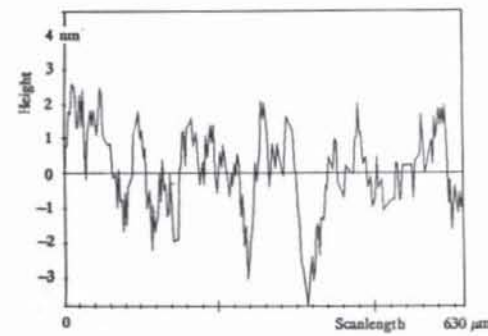
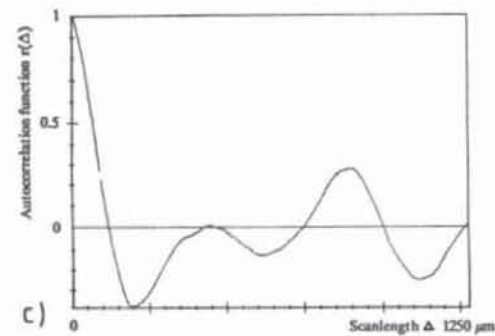
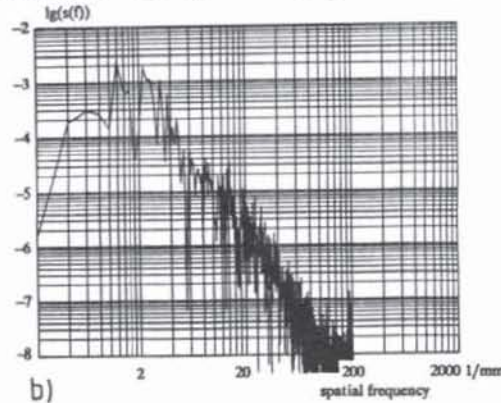
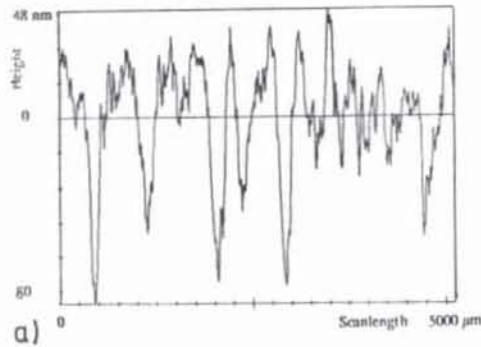


Fig. 6

Fig. 7

local slopes and edges are extremely difficult to be profiled. Double-pass compensates for scattering angles which pass the entrance pupil of the scanning lens. But stronger scattering passing outside of the pupil causes fatal break down of the signal.

Fig. 6 shows an example of a successfully scanned surface with mean roughness $R_a = 20\text{nm}$ and relatively large local slopes. The profile of a very smooth zerodur sample (Fig.7) has a rms-value of less than $1\ \text{nm}$ after high-pass filtering with a cutoff wavelength of $100\ \mu\text{m}$.

3.2 Photometric balance profilometer This principle of height detection originated from focussing methods^{8,9}. Fig.8a shows our realization with laser^{10,11} and double-pass for compensation of local slopes.

The gaussian beam of a He-Ne-Laser is focussed onto the rough surface by the profiling lens O3. Two photodetectors DET1 and DET2 are positioned intra- and extrafocal with respect to the secondary foci F1 and F2 of the scanning focus on the surface. The detectors DET1 and DET2 are adjusted so as to receive exactly the same light fluxes when the surface is exactly in focus. The normalized difference signal $s = (I_1 - I_2) / (I_1 + I_2)$ is zero in this case. If however the surface is scanned, the signal s is a measure of the profile height. The height $y(x)$ is calculated for every data point with a calibration curve stored in the computer as a polynomial of third order as the best fit within the calibration range K of the characteristic curve of Fig. 8b.

As in the case of the heterodyne profilometer large errors are produced, when large local slopes, curvatures, edges, or structures within the width of the focus occur. On surfaces not too rough our photometric profilometer, when operating in this simple mode, has a resolution of $R_q \approx 2-5\text{nm}$. In order to improve the dynamic range, a automatic focus control, AFC, is incorporated and the vertical position of the scanning lens O3, when it is exactly in focus, is taken as the relative profile height. In this operation mode a range as large as 1:5000 can be realized. Fig.9 shows a typical result of our high resolution photo-

(Fig.5a). The depths of the measured grooves agree well with the values given by the manufacturer. Heterodyne profilometry is well suited to measure geometrical structures without large local slopes and sharp discontinuities.

Fig.6: Hand-polished metal surface with $R_q=20\text{nm}$
 a) high-pass filtered profile cut-off wavelength $L=800\ \mu\text{m}$
 b) normalized spectral density function (periodogram)
 c) autocorrelation function of the filtered profile
 Fig. 7: Smooth zerodur glass-ceramic with sub-nm resolution; cut-off wavelength $L_c = 100\ \mu\text{m}$

However surfaces with irregularities within the spot width or with large

metric profilometer. It is a dark and very hard SiC ceramic gasket ring, very costly to be profiled with stylus instruments because of a possible damage of the diamond tip.

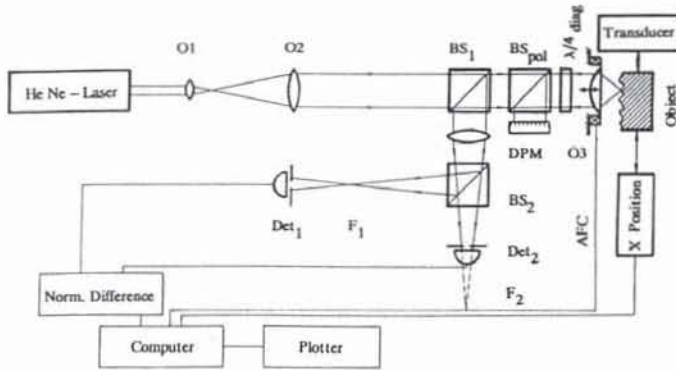


Fig. 8a: arrangement of the photometric profilometer

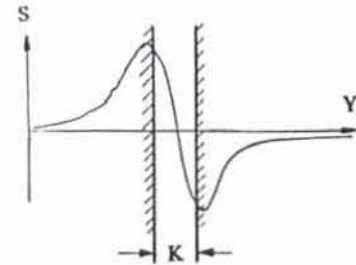


Fig. 8b: kalibration range

Commercially available probes based on the photometric principle but using compact-disk technology have resolution and measuring range given in table 1.

Fig. 10a shows a 3-D scan of 64x64 sample points of a end-milled metal surface of nominal mean roughness $R = 1,6 \mu\text{m}$ and Fig. 10b a graphical representation of the 2-dimensional spectral density function.

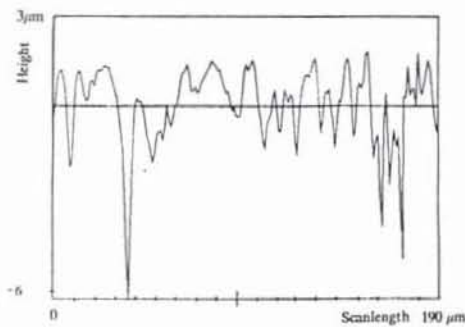


Fig.9 SiC-ceramic

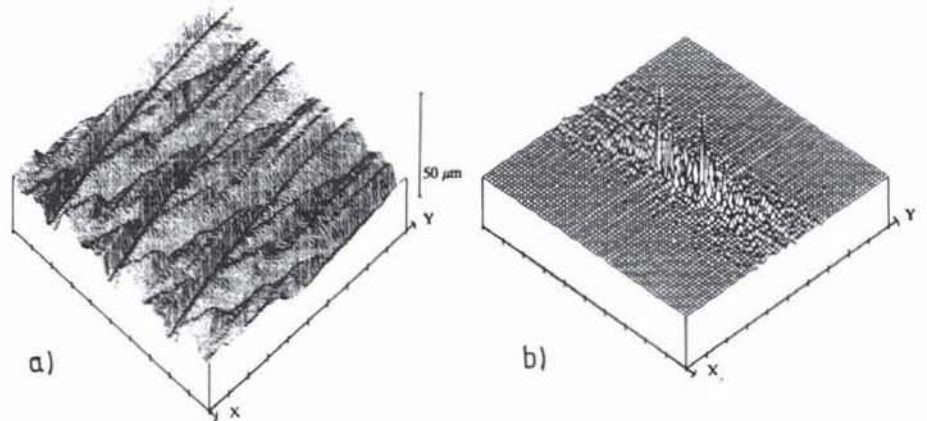


Fig.10 End-milled surface

3.3 Compensation of local slopes by double-pass operation

Fig 11a shows the effect of local surface slopes in a photometric profilometer. A slope angel of α leads to geometric vignetting. A local defect can produce large "spikes" in the measured surface and fine surfaces are measured with incorrect roughness values.

	measuring range	resolution	lateral resolution 4)	slope compensation	technologie
UBM 16 ^{a)}	500 μm^1 50 μm^1	100 nm 10 nm ²⁾	500 lines/mm 500 lines/mm	-- --	CD CD
RM 600 ^{b)}	300 μm^1 30 μm^1 3 μm^1	200 nm ¹⁾ 20 nm ¹⁾ 2 nm ^{1), 2)}	500 lines/mm	--	CD
ITOC ^{c)} Phot-Prof	20 μm^3	4 nm ²⁾	1200 lines/mm	double-pass ⁵⁾	laboratory set-up
ITOC ^{c)} Het-Prof	20 μm^3	< 0,5 nm ²⁾	1200 lines/mm	double-pass ⁵⁾	laboratory set up

- a) U. Breitmeier Meßtechnik GmbH, Stuttgart
- b) G. Rodenstock, München
- c) Institut für technische Optik, Stuttgart
- 1) data given by the manufacturer
- 2) on relatively smooth surfaces
- 3) limited presently by a piezo-translator of the AFC
- 4) spacial frequency with transfer factor 0,5
- 5) single-pass optionally

table 1.

Fig. 11b shows the principle of the double-pass system C. The double-pass mirror in the back focal plane on the scanning objective reflects the deviated and scattered beams back into the back focal plane of the surface from where the beams are reflected backwards exactly in their original direction (usually the beam axis). The ideal double-pass mirror would be a phase conjugated mirror which would compensate also for local curvature. In Fig. 12a polarizing beamsplitter and a $\lambda/4$ plate in diagonal orientation show the same effect but maintain the full aperture. By rotation of the phase-plates by an angle of $\pi/4$ the operation can easily be changed to single pass, for surfaces of very low reflection.

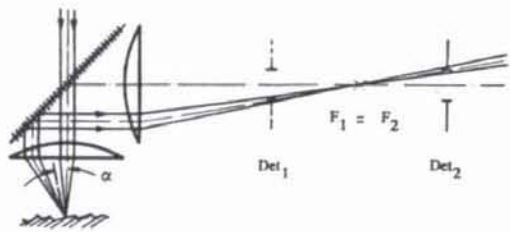


Fig. 11a

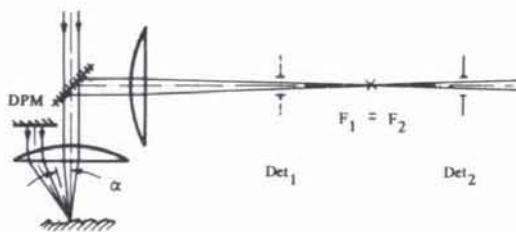


Fig. 11b

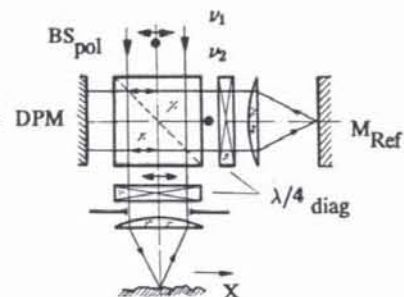


Fig. 12

4. Roughness measurement with the interference microscope and automated fringe analysis phase-shifting interferometry

For highly polished surfaces with roughness values $R_q < 25$ nm interference microscopy such as a Nomarski differential interference contrast is a well established method of roughness estimation. Modern methods of automated fringe analysis by phase shifting interferometry^{12,13,14}, can be used for smooth surfaces. Fig. 13 shows our modified Linnik Interference-microscope. Thus, microscope objectives with very high numerical aperture can be used and very high lateral resolution can be obtained.

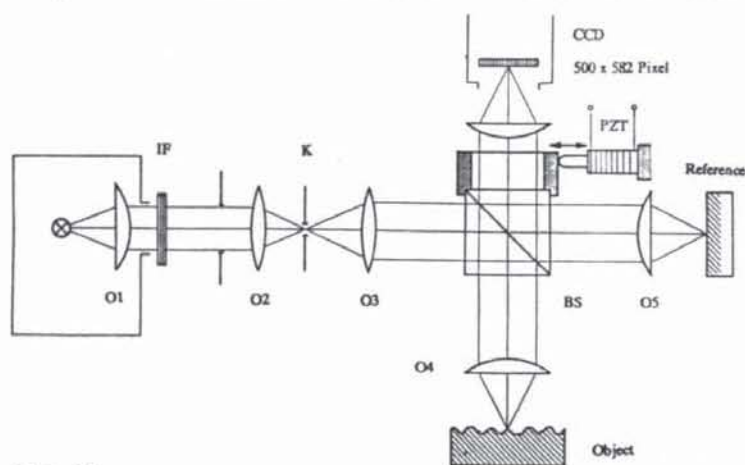


Fig. 13

White light of a incandescent lamp is used. A white light fringe system as close as possible to the central or achromatic fringe is adjusted. Immediately before the phase steps are introduced, interference filters are inserted into the illuminating beam. For automatic fringe analysis the fringes are shifted by the piezo translator acting onto the beamsplitter BS.

Three intensity distributions¹⁵ are sampled by a 500x582 pixel image sensor and digitized by a image processing card in a AT-PC computer. The interference phases $\phi(x,y)$ of each pixel within a preselected frame and $h(x,y)$ are calculated.

$$h(x,y) = \frac{\lambda}{4\pi} \phi(x,y) \quad (6)$$

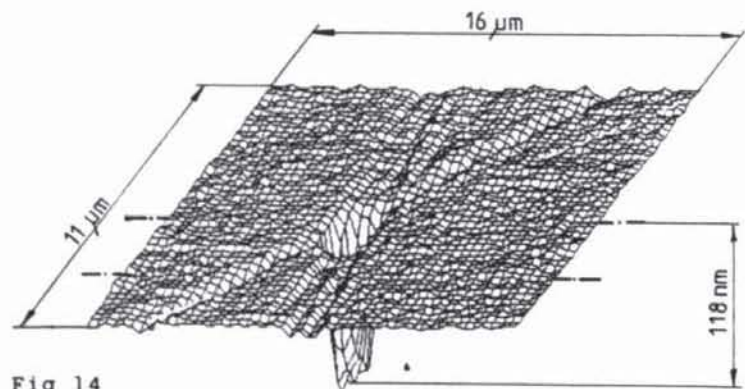


Fig 14

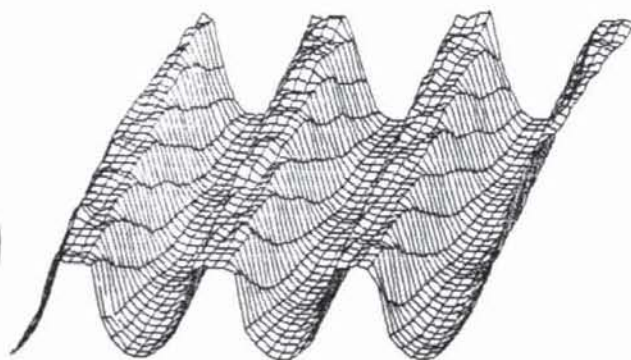


Fig 15

However, as in the case of heterodyne- and photometric profilometry, surfaces with large local slopes, edges, and irregularities within the point spread are difficult to be mapped. Phase- and localization errors arise even for surfaces with height variations not exceeding the $\lambda/2$ - intervall. Fig. 14 shows a polished germanium surface with scratches and a defect. 50x50 pixels were used for an object field of $16 \mu\text{m} \times 11 \mu\text{m}$. Fig. 15 shows an example of a coarser geometrical structure. It is a grid-like topography from a semiconductor chip with maximal vertical height difference of 238 nm. For height variations larger than $\lambda/2$ two wavelength phase-shifting methodes can be used.

5. Integral roughness measurement

While photometric and interferometric methods are mapping the surface in one or two scanning directions, integral or area methods use the scattered radiation from a finite area of the surface and supply statistical characteristics of the surface.

5.1 Coherent scattering It is well known that coherent scattering in the far field or in the image plane of a lens is not well suited for roughness $R_q \geq \lambda/5$. In¹⁴ some calculations of the scattered intensity for the small angle range in the fourier-plane of a scattering detector are shown as a function of vertical roughness R_q and of the horizontal roughness parameter correlation length w_x . For the R_q -range interesting for industrial produced metallic surfaces, the intensity is shown to be a function of both R_q and w_x . Thus there is apparently no way to separate and to measure both parameters independently.

5.2 White light speckle contrast The same situation is found for coherent speckle contrast measurement (e.g.^{16,17}). As it is well known the speckle contrast C

$$C = \frac{\sigma_I}{\langle I \rangle} = \left[\frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1 \right]^{1/2} \quad (7)$$

where $\langle I \rangle$ is the mean intensity over a sample of intensity measurements I and σ_I is the standard deviation, saturates to a value $C \approx 1$ for $R_q \geq \lambda/5$, provided the number of independent scattering regions on the surface is large enough (see Fig. 18a). As the optical fields near the detector are nearly gaussian, experiments involving higher order moments are also of the restricted value.

Speckle measurements with white light^{18,19} have shown that the contrast curves C (Fig. 19b), after reaching a maximum contrast for $R_q \approx \lambda/5$ fall down again. However, due to nonmonotonic behavior of measured curves and due to the interdependency of horizontal and vertical roughness parameters no practical application of direct white light speckle measurement has become known so far.

In^{20,21} the nonmonotonic behavior of the measured contrast curves could be overcome by a incoherent superposition of a uniform intensity to the speckle field near the detector. Fig. 16 shows a schematic arrangement for a contrast sensor. White light speckle are scanned by the pinhole of a photodetector or by a detector array in a defocussed image plane x' . In²⁰ a monotonic falling curve of contrast as a function of R_q was obtained by an incoherent superposition of light from the same beam by beamsplitter BS, attenuator BA, and mirror M.

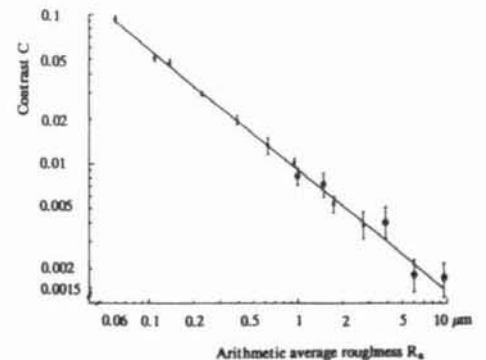
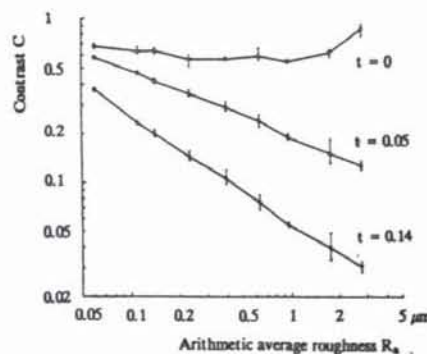
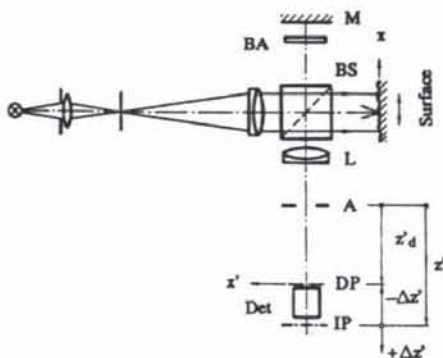


Fig. 16

Fig. 17

Fig. 18

Fig. 17 shows the dependence of the contrast C with the ratio t of constant intensity to speckle intensity for $R_q \hat{=} \sigma_h = 0$:

$$t = \frac{I_{\text{const}}}{I_S | \sigma_h = 0} \quad (8)$$

It can be seen that rather small values of incoherently superposed light (straylight!) can affect drastically the contrast curves. Fig.18 shows the contrast C measured from ground and planed standard specimens with roughness $0.06\mu\text{m} < R_a < 10\mu\text{m}$. Due to a ratio $t = 1.4$ the measuring range $0.06\mu\text{m} < R_q < 10\mu\text{m}$ is rather large and exceeds the ranges of conventional speckle contrast methodes.

5.3 White light elimination method In²² we describe a new method. Vertical descriptor R_q and horizontal descriptor autocorrelation width w_a are determined as independent values.

The same sensor as in Fig.15 is used, however, contrast C_0 from a measurement with $t \approx 0$ and in addition C_m from a measurement with $t = 1.4$ is calculated. Fig.19b and 19c show curves of speckle contrast from a unpublished theoretical model using a special integration method over interlaced autocorrelation- and coherence-facettes on the surface involving a triangular autocorrelation function of width w_a . The end formula is given by eqn.(4) of²² as a function of σ_h , w_a and t and of

- k = wavenumber = $2\pi/\lambda$, λ = wavelengs
- W = width of the spectral density of illumination
- w_b = width of the geometric point spread funktion
- w_c = width of the spatial coherence,

The contrasts have been found to be in good agreement with experimental values (except for parameter combination; where $\sigma_h > 5\mu\text{m}$, $t < 0.05$). The curves give an excellent overview over contrast behavior as a function of all the parameters listed above.

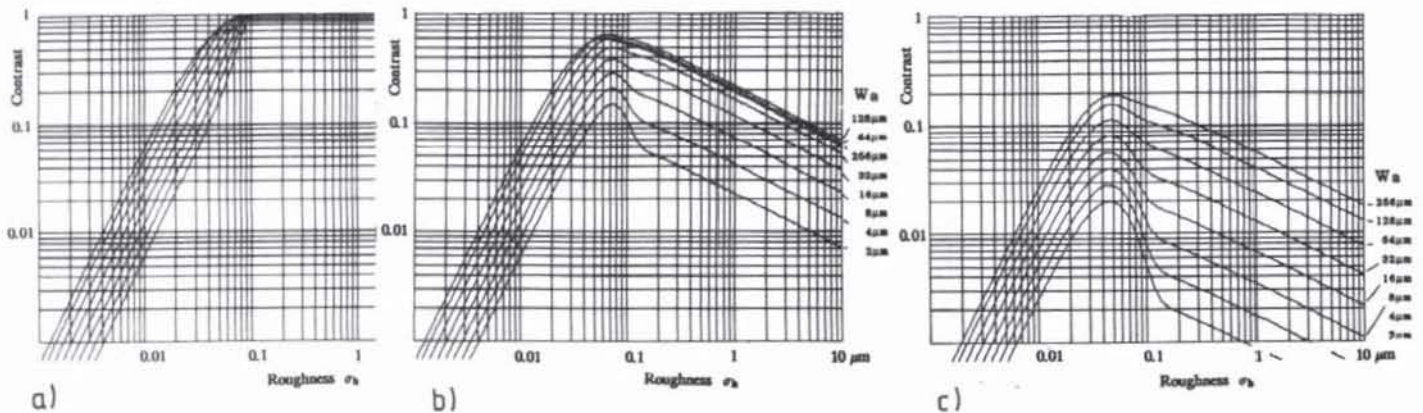


Fig.19

We see from Fig.19 that for $R_q \hat{=} \sigma_h > 0.15 \mu\text{m}$ all curves in b) and c) are parallel to each other. Thus the ratio $c = C_0/C_m$ in this theory is independent from R_q . We find from eqn.(4) of²²

$$w_a = \frac{(2w_b - w_c) w_c}{(c-1) w_b/t + w_b - w_c} \quad (9)$$

w_b , w_c , and t are known parameters of instrument and the light, and c is given by measuring C_0 and C_m as described above. Thus, the autocorrelation width w_a can be determined independently of σ_h . For $\sigma_h > 0.15\mu\text{m}$ eqn.(4) of²² reduces with good approximation to

$$C^2 = \nu_1 / \left[(1 + 16W^2\sigma_h^2)^{1/2} ((\nu_1 + w_b/w_a)t + w_b/w_a)^2 \right] \quad (10)$$

$$\nu_1 = (2 w_b/w_a - w_c/w_a)(w_c/w_a - 1) \quad (11)$$

For w_a known by (9) the uncertainty of the determination is resolved and (10) can be inverted for the unknown σ_h :

$$\sigma_h = (1/4 W) \left[\frac{\nu_1^2}{C^4 \left[(\nu_1 + w_b/w_a)t + w_b/w_a \right]^4} - 1 \right]^{1/2} \hat{=} R_q \quad (12)$$

The root-mean-square roughness $R_q \hat{=} \sigma_h$ can thus be calculated in quasi real time by the same PC- computer as used for sampling the intensity values.

Practical measurements on surfaces with known roughness- and autocorrelation-values have shown very good results for $0,15\mu\text{m} < R_q < 2\mu\text{m}$ and even, though not predicted by theory, for smaller R_q down to $R_q=0,03\mu\text{m}$. Fig. 20 shows results of this procedure for surfaces of different manufacturing processes as grinding, polishing and planing.

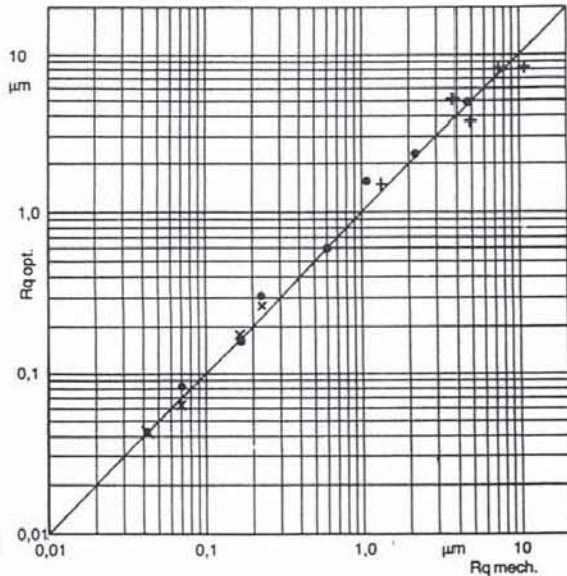


Fig. 20

In order to improve the method for surfaces with $R_q > 2\mu\text{m}$ and to extend the measuring range up to $R_q=10\mu\text{m}$, correction coefficients have been introduced in (12). Further investigation on this coefficients and their justification as well as on the measured values of w_a will be necessary.

A sensor based on this method with an array of 1728 diodes has a measuring time of about 3 ms for each of the two contrast measurements. The light source can be a small 10 W incandescent lamp and the free measuring distance is about 70 mm.

Thus apart from the mapping methods of profilometry and phase-shift interference microscopy, new integral roughness sensors of the type described above could be very promising for industrial on-line applications.

Acknowledgments. Part of the work leading to this paper was supported by Deutsche Forschungsgemeinschaft under DFG Ti 119/4-1 and /4-2.

6. References

- 1 J.M.Bennet, Comparison of techniques for measuring the roughness of optical surfaces, Opt. Engineering 24, 380- 387 (1985).
- 2 T.V.Vorburger, Measurement of roughness of very smooth surfaces, Ann. CIRP 36 (1987)
- 3 J.S. Bendat and A.G. Piersol, Random Data, Wiley, New York (1971).
- 4 A. Papoulis, Propability, Random variables and Stochastic Process, Mc.Graw Hill (1965)
- 5 E.L.Church, Fractal surface finish, Applied Optics 27, 1518- 1526 (1988).
- 6 G.E.Sommagren, Optical heterodyne profilometry, Applied Optics 20, 610- 618 (1981).
- 7 H. K. Wickramasinghe, Differential laser heterodyne micrometrology, Opt. Engineering 24, 926- 929 (1985).
- 8 O.Dupuy, Methode de pointes longitudinaux de grande precision, Revue d'Optique 43, 217- 244 and 282- 306 (1964).
- 9 J.Simon, New noncontating devices for measuring small microdisplacements, Applied Optics 9, 2337- 2340 (1970).
- 10 K. Leonhardt, K.-H. Rippert und H.J. Tiziani, Verfahren zur optischen Rauheitsmessung und Mikroprofilometrie, PTB-Bericht, PTB-Opt. 19, 80- 111 (1985).
- 11 K. Leonhardt, K.-H. Rippert und H.J. Tiziani, Optische Mikroprofilometrie und Rauheitsmessung, Technisches Messen, tm 54, 243-252 (1987).
- 12 J.H. Bruning et al., Digital wavefront measuring interferometer for testing optical surfaces and lenses, Applied Optics 13, 2693- 2703 (1974).
- 13 J.C. Wyant et al., An optical profilometer, ASLE Transactions 27, 101-113 (1984).
- 14 H.J.Tiziani, Rechnerunterst.Laser-Meßtechnik, Technisches Messen, tm 54, 221- 230 (1987)
- 15 B. Dörband, Die 3-Interferogramm-Methode..., OPTIK 60, 161-174 (1982)
- 16 J.C. Dainty, Editor, Laser speckle and related phenomena, Springer, Berlin (1975).
- 17 Asakura, T., Surface roughness measurement in Erf, K. Ed., Speckle metrologie Academic Press, New York (1978).
- 18 R. A. Sprague, Surface roughness measurement using white light speckle, Applied Optics 11, 2811-2816 (1972).
- 19 K.Nakagawa und T.Asakura , Contrast of white light speckle patterns at a defocussed image plane, Applied Optics 18, 3725- 3728 (1979).
- 20 K. Leonhardt und B. Pfister, Ein neues optisches Verfahren zur schnellen, berührungslosen Rauheitsmessung, OPTIK 58, 297-319 (1981).
- 21 K. Leonhardt und H.J. Tiziani, Removing ambiquities in surface roughness measurement, OPTICA ACTA 29, 493- 499 (1982).
- 22 K. Leonhardt, E. Kaufmann and H.J. Tiziani, Determination of average roughness and profile autocorrelation width of metallic surfaces with a white light sensor, Optics communic. 51, 363- 366 (1984).