Journal of Visualization, Vol. 3, No. 1 (2000) 63-70

Visualization of the Surface Roughness by X-ray Scanning

Protopopov, V. V.* and Danilov, V. V.*

* The Institute of Physics and Technology, Russian Academy of Sciences, 117218 Moscow, Nakhimovskii Pr. 34, Russia.
e-mail: proto@labdif.crystal.msk.su

Received 14 June 1999.
Revised 17 August 1999.

***Abstract*:** In the x-ray region the reflectivity of a superpolished surface strongly depends on its roughness. This effect may be used to obtain a two-dimensional map of the roughness spatial distribution for flat surfaces with an average roughness height of the order of one nanometer or less. The method described in the paper lies in the illumination of the sample by a highly collimated x-ray beam, and a linear one-dimensional scanning of the sample with simultaneous registration of the specular component of the reflected beam by multielement linear detector. This method may be used to monitor the surface quality of silicon semiconductor wafers, computer hard disks, x-ray and laser mirror substrates etc.

***Keywords*:** roughness, x-ray, reflection, scanning, linear detector.

1. Introduction

The progress in microelectronics, computer components technology, x-ray and laser optics demands creation of supersmooth flat surfaces with a nanometer scale roughness, such as silicon semiconductor wafers, computer hard disks, x-ray and laser mirror substrates. These practical tasks stimulate the search for efficient methods for the monitoring of the roughness of surfaces with an area of $10^4 mm^2$ and larger. Today, the routine method for studying the surfaces with nanoscale resolution is scanning probe microscopy. But the extremely small area (approximately $10 \times 10 \mu m^2$) covered by a single such measurement makes the monitoring of the whole sample in an acceptable time interval virtually impossible. Moreover, the sensitive element of a scanning probe microscope touches the surface, that is in many cases unacceptable. Therefore, new methods of remote sensing are welcomed, that radically reduce the duration of the measurements, preserving at the same time the roughness resolution at a level of several angstroms and a spatial resolution of the order of 1mm.

One of the possible solutions to this problem is the measuring of the intensity of x-rays, reflected from the tested surface. In some interval of angles out of the region of total external reflection the reflectivity falls rapidly with roughness. This effect, in more details discussed in the next section, may be used to obtain a two-dimensional map of the roughness spatial distribution for flat surfaces with an average roughness height of the order of one nanometer or less. Such a map may be presented in a visualized form by means of computer programs, thus making the usually invisible defects of a tested surface visible.

2. Reflection of X-rays from Rough Surfaces

While the wavelength of the visible light lies in the interval of $0.4 \div 0.6$ mkm, the wavelength of the x-rays is typically 10^{-4} mkm (1Å) or less. Therefore, the surface roughness with the average height of the order of 10^{-3} mkm (nano-scale roughness), to which the visible light is practically insensitive, strongly influences the reflectivity of the x-rays. The x-rays obey the same physical laws, as a visible light does. But, unlike the visible region of wavelengths, in the x-ray region the real part of the refractive index is less than unity, and the absorption plays a

significant role.

According to the Fresnel law (see, for example, Born and Wolf, 1968), the coefficient of intensity reflection from an ideally smooth surface for a nonpolarized plane wave is equal to

$$R_0(\theta) = \frac{1}{2}\left[\left|r_s(\theta)\right|^2 + \left|r_p(\theta)\right|^2\right];$$

$$r_s(\theta) = \frac{\sin\theta - \sqrt{\sin^2\theta + \chi}}{\sin\theta + \sqrt{\sin^2\theta + \chi}}; \qquad (1)$$

$$r_p(\theta) = \frac{(1+\chi)\sin\theta - \sqrt{\sin^2\theta + \chi}}{(1+\chi)\sin\theta + \sqrt{\sin^2\theta + \chi}};$$

$$\chi = -\chi_r + i\chi_i.$$

Here, θ is the angle between the surface and the wave vector, usually referred to as a grazing angle, χ_r and χ_i are the real and the imaginary parts of the dielectric susceptibility respectively, and λ is the wavelength. In the x-ray region the real part of the dielectric susceptibility is a very small quantity. For instance, for $\lambda = 1.54$Å (CuK_α radiation) χ_r is of the order of $10^{-4} \div 10^{-5}$, and χ_i is of the order of $10^{-8} \div 10^{-5}$. Therefore, x rays may be effectively reflected only in the region of rather small grazing angles, not exceeding $0.2° \div 0.3°$ for typical materials. This is the so-called total external reflection region.

The roughness decreases the reflection coefficient exponentially (Isakovich, 1952):

$$R(\theta, \sigma) = R_0(\theta) \cdot \exp\left[-\sigma^2 f(\theta)\right], \qquad (2)$$

where σ is the root mean square roughness height, and $f(\theta)$ is the known function of a grazing angle θ (Névot and Croce, 1980):

$$f(\theta) = \left(\frac{4\pi}{\lambda}\right)^2 \sin\theta \cdot \mathrm{Re}\sqrt{\sin^2\theta + \chi}. \qquad (3)$$

The angular dependence of the reflection coefficient is presented in Fig.1 for two values of σ. The grazing angle, at which the reflection coefficient falls abruptly, is known as the critical angle θ_c. Clearly, it is possible to evaluate σ by measuring the intensity I of the reflected wave at different points of the surface at grazing angles larger than critical one:

$$\sigma = \sqrt{c_1 + c_2 \ln I}, \qquad (4)$$

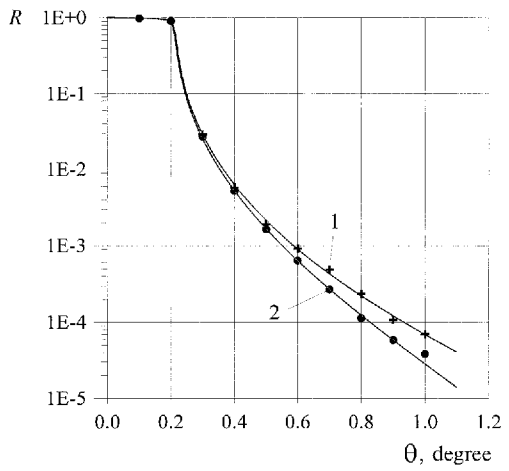

Fig. 1. Experimental data (dots and crosses) and theoretical fits (solid curves) of the reflection coefficients for two SiO_2 samples with different roughness:
(1) – $\sigma = 5.8$Å, $\theta_c = 0.213°$, $\chi_i = 2.55 \cdot 10^{-7}$; (2) – $\sigma = 8.7$Å, $\theta_c = 0.210°$, $\chi_i = 2.09 \cdot 10^{-7}$. CuKa radiation ($\lambda = 1.54$Å).

where c_1 and c_2 are parameters, depending on the grazing angle and optical constants of the sample. Optical constants do not vary along the sample surface, and nor does the grazing angle if the surface is plane. Therefore, the roughness spatial distribution of a plane surface may be determined unambiguously. But, if the surface is not a plane one, the grazing angle may vary from one point to another, and, consequently, the value of σ, calculated by (4) with constant c_1 and c_2, may differ from its real value.

Suppose we have to resolve two spatial regions with roughnesses σ_1 and σ_2. The larger the difference $R(\theta,\sigma_1)-R(\theta,\sigma_2)$, the better the conditions for separating these two regions. From this point of view, one should choose as large grazing angle θ as possible. On the other hand, the larger the grazing angle, the lower the intensity of the reflected x-ray beam, and, hence, the greater role of the random fluctuations of the detector signal. Thus, there must be an optimal value of grazing angle θ. Let the number of photons in the incident x-ray beam be equal to N. Then the difference between the signals from the regions with roughnesses σ_1 and σ_2 is equal to

$$N\left[R(\theta,\sigma_1)-R(\theta,\sigma_2)\right]\approx N\frac{\partial R(\theta,\sigma)}{\partial\sigma}\Delta\sigma, \qquad (5)$$

$$\Delta\sigma=\sigma_1-\sigma_2.$$

The amplitude of random fluctuations in the detector signal is proportional to the sum of the root mean square deviation of the photon number in the reflected beam and the equivalent detector noise η:

$$\sqrt{NR(\theta,\sigma)}+\eta. \qquad (6)$$

It is possible to introduce a parameter q, equivalent to the signal-to-noise ratio:

$$q=\frac{N\Delta\sigma\cdot\partial R(\theta,\sigma)/\partial\sigma}{\sqrt{NR(\theta,\sigma)+\eta}}$$

$$=8\sigma\Delta\sigma k^2\cdot\frac{\sin\theta\sqrt{\sin^2\theta+\chi}\cdot NR(\theta,\sigma)}{\sqrt{NR(\theta,\sigma)+\eta}}. \qquad (7)$$

The function $q(\theta)$ is presented graphically in Fig. 2, from which it follows that the optimal grazing angle is equal to approximately 0.3° for Si. Practically the same value is optimal for other light materials, such as SiO_2 and Al. For heavy materials, such as Mo and W, the optimal grazing angle is larger: 0.5°÷0.6°. For the samples with roughness levels greater than 10Å, the optimal value of the grazing angle must be strictly observed for the sake of a high signal-to-noise ratio. For the samples of higher surface quality, with roughness of the order of 5Å and less, this condition is not so restrictive, because greater grazing angles, for instance, 0.5°÷0.6° instead of 0.3°, give practically the same signal-to-noise ratio.

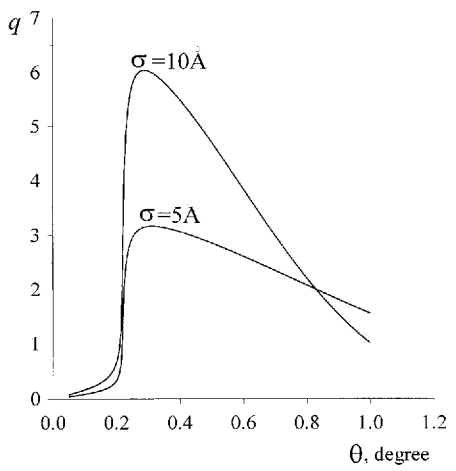

Fig. 2. The angular dependence of the signal-to-noise ratio for Si-substrate. $N=2\cdot10^6$, $\eta=10$, $\Delta\sigma=1$Å. CuK_α radiation ($\lambda=1.54$Å).

3. The Optical Scheme of the Method

The idea of this method can be clearly understood from Fig. 3 (Protopopov et al., 1998). A parallel x-ray beam comes to a sample at the grazing angle θ, while the sample is moving in the direction perpendicular to the incidence plane. The sample dimension D, the grazing angle θ, and the beam width in a vertical direction d, satisfy the straightforward relation $d = D \sin \theta$. If the surface of a sample is a plane, the reflected beam is also parallel, so that the information about every surface element δ will be relayed to a detector in a narrow parallel beam of the width $\delta \cdot \sin \theta$ in the vertical plane. Thus, if we want the spatial resolution in the sample plane to be δ, the detector spatial resolution has to be equal to $\varepsilon = \delta \sin \theta$. For many practical purposes, spatial resolution of $\delta = 1$mm is sufficient. The optimal value of a grazing angle, as it is shown in the section 2, lies in the interval $0.3° \div 0.6°$. Therefore, the detector spatial resolution should be equal to $\varepsilon = 5 \div 10$ mkm. In our experiments the detector spatial resolution was equal to 12 mkm, that was close enough to the desired value.

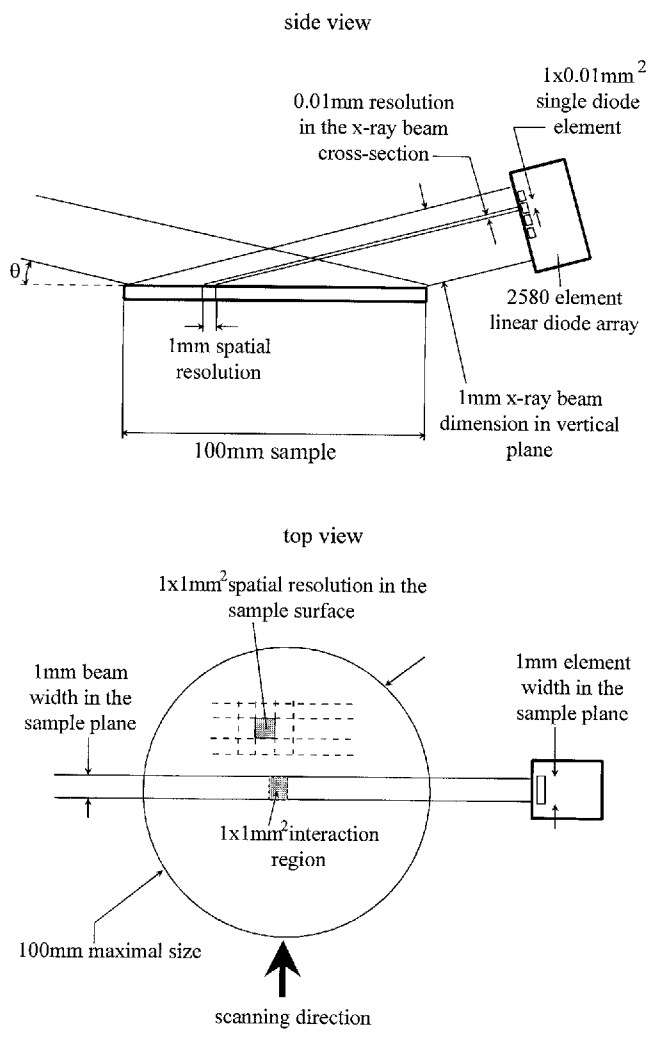

Fig. 3. The optical scheme of the method.

Suppose a sample dimension is 100×100mm^2. Then the number of resolution elements in the direction of the x-ray beam is equal to 100. In order to have the same resolution element number along the orthogonal axis, we have to limit the beam width in the sample plane to 1mm, which also matches the single detector element width in this direction. Thus, the total number of pixels in the image will be 100×100.

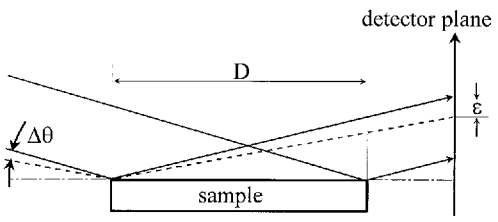

Fig. 4. The influence of a beam divergence.

There are several reasons why the incident x-ray beam coming to the sample must be highly collimated in the vertical plane. First of all, the grazing angles at every point of the sample surface must be identical in order to produce equal signals from the surface elements with equal roughness levels. Secondly, a parallel beam does not introduce a distortion in the image. In addition, the divergence of the incident beam affects the resolution negatively. The sketch in Fig. 4 explains the situation. Suppose we want the roughness resolution of the order of 1÷2Å not to be affected by the divergence of the incident beam. Then, it follows from Fig. 1 that for grazing angles in the region of 0.5°÷0.6°, the limitation $\Delta\theta < 0.02°\div0.03°$ must take place. On the other hand, in order for the divergence to not affect the spatial resolution, the relation $D\cdot\Delta\theta < \varepsilon$ must take place, where ε is the vertical size of the detector single element. In our case, $D \approx 100$mm and $\varepsilon \approx 0.01$mm, so that $\Delta\theta < 10^{-4}$ radian $\approx 0.006°$. Thus, the spatial resolution limitation is far more restrictive than that of the roughness resolution. Therefore the collimating optics, designed to limit the divergence of the x-ray beam to a value of 10^{-4} radian, should be installed between an x-ray tube and the sample.

4. Experimental Results

To show the feasibility of the method, described above, we used both specially prepared test objects and real samples, such as high power laser mirrors, hard disks etc. One of the test objects, designed for the investigation of a spatial resolution, was prepared on a superpolished flat quartz substrate of 70mm in diameter with an average surface roughness of 4Å and a maximal surface profile distortion of less than $4\cdot10^{-5}$ radians. The flatness of the substrate guaranteed that the surface profile distortion would not affect the image quality. Six tungsten stripes 6, 4, 3, 2, 1.5 and 1mm in width and about 300Å thick were deposited on its surface as shown in Fig. 5(a).

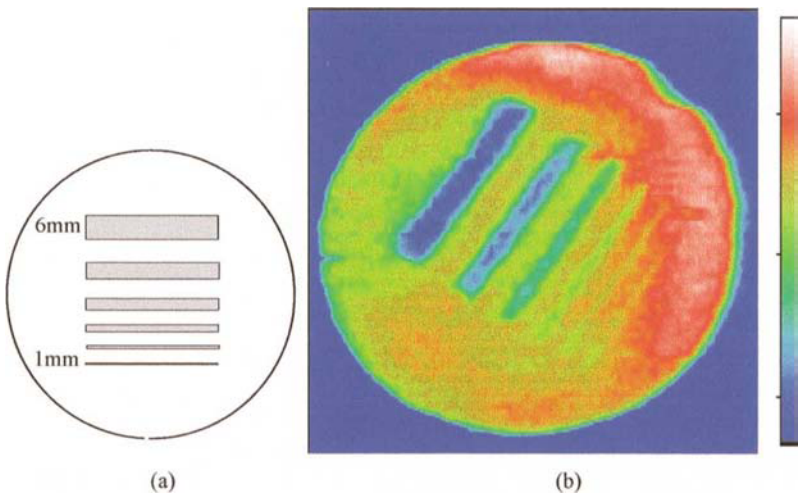

Fig. 5. The scheme (a) and the image (b) of a test object.

Due to the stochastic character of the deposition process, the roughness of the stripes was essentially larger than that of the substrate: approximately 8Å. Afterwards, a relatively thin tungsten layer of approximately 30Å was deposited onto the entire surface. The roughness of this layer in the regions outside the stripes was expected to be less than the roughness of the stripes. The x-ray measurements gave a value even less than that of the substrate: ≈ 2Å. Thus, the difference between the roughness levels of the stripes and the remaining surface was about 6Å.

The image of this object is presented in Fig. 5(b). Pseudo colors of this image, generated by a computer program P4D_SPM, are determined by the intensity of the reflected x-ray beam according to a color scale bar to the right of the picture, where the upper lying colors correspond to a larger reflected intensity. Therefore, according to (2), darker green and blue regions of the image correspond to a large roughness. One can see that all six stripes are resolved, though the contrast of the first three widest stripes is well above the noise level, while that of the others is poorer. The stripes in the image are of a dark tone, i.e., the reflection from these regions is less than that from the others. This proves that the nature of the image contrast is the roughness variation and not the variation of the average thickness of the layers.

Using transformation (4) and computer-generated pseudo colors, it is possible to present the image in terms of a root mean square roughness height. Additional visual effect may be obtained by portraying the data in a 3D form, as is shown in Fig. 6. Here not only the color scale is calibrated in r.m.s. roughness height (nanometers in this particular case), but the vertical axis of the image also displays the value of the roughness.

The next example shows, how defects of computer hard disks may be effectively visualized by x-ray scanning. Figure 7 presents the image of a hard disk of 100mm in diameter. The color scheme is the same as that of Fig. 5. Arrows point to local variations of surface roughness, while the frames show locally convex or concave regions, which may be readily selected, taking into consideration alteration of colors or deformation of the object border. The green inner ring shows the magnetic head landing zone with a higher roughness.

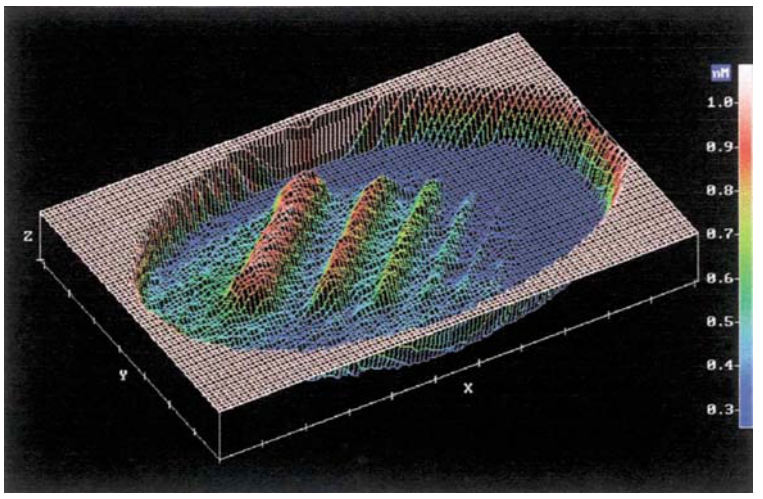

Fig. 6. A computer-generated 3D view of the roughness spatial distribution.

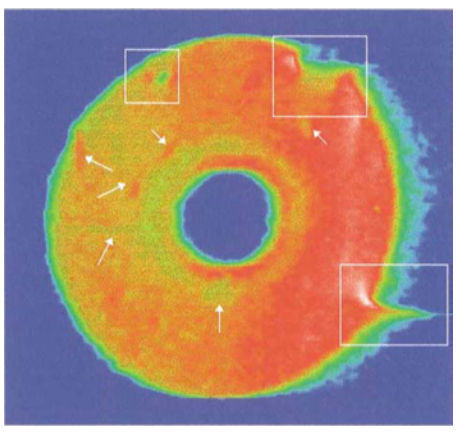

Fig. 7. An image of a hard disk defects.

5. Image Refinement

Two different processes form the image in the x-ray scanning method: mechanical movement along one axis and a linear detector array along another. Therefore, different resolution limitations will take place along different axes. The width of the x-ray beam and the speed of the mechanical movement, if needed, may be lowered in order to minimize their influence on the resolution. As for the detector, there is quite a different situation: its physical properties cannot be changed so easily. For instance, diffuse scattering of x-rays on surface roughness, high energy photoelectrons produced by x-rays, penetrating the detector sensitive elements, electrical losses and time delay during the signal transfer from sensitive elements to counting circuits, and some other reasons, limiting detector spatial resolution, cannot be eliminated. Thus, any image deterioration in our case is primarily due to detector limitations. Therefore, the discrete signal I, obtained in every line of an image, may be expressed in terms of a system of linear equations:

$$I(\chi_i) = \sum_j h(y_j - x_i) V(y_j) + n_i, \tag{8}$$

where V is an unknown unperturbed line signal, n is an additive detector noise, and h is a known point-spread function of a detector. In our case a good analytical approximation for h was found to be a Lorentzian function:

$$h(x) = \left[1 + \left(\frac{x}{\omega}\right)^2\right]^{-1}, \tag{9}$$

with ω equal to 2 sensitive elements of the detector.

In principle, if n_i are known values, it is possible to find V, solving (8) directly with the help of any method for solving systems of linear equations. But the random noise signal is unknown. Moreover, the number of equations is often so large (more than 100) that it is impossible to solve system (8) directly, retaining necessary accuracy. Nevertheless, it is possible to evaluate V approximately, applying a least squares method, that is, finding such values of V, that deliver a minimum to a functional

$$F(V) = \sum_i \left[I(x_i) - \sum_j h(y_j - x_i) V(y_j) \right]^2. \tag{10}$$

There are several standard routines for solving this problem, for example, the method of conjugated gradients (Gill et al., 1981).

Using the aforementioned methods, we have developed a computer code, and applied it to the images of the roughness spatial distribution. The result is shown in Fig. 8(b), comparing it with the initial image, discussed in previous section. One can easily see that refinement procedure enhances the contrast of defects, especially of point-like ones. For instance, the refined image clearly show that the defect, marked by the arrow, has a form of a

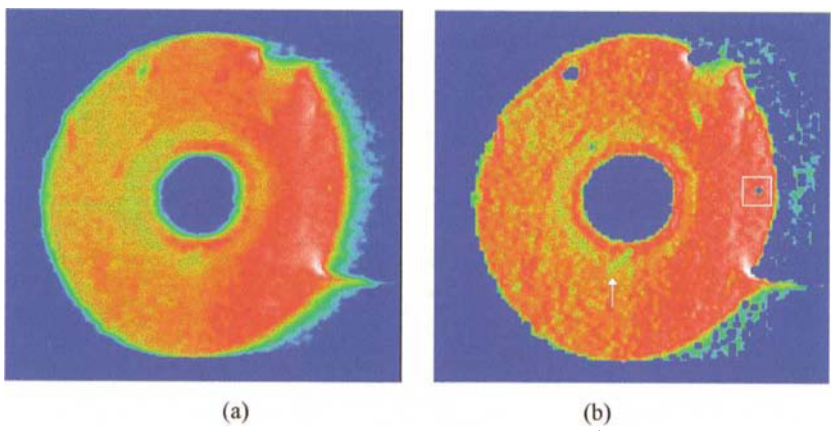

Fig. 8. An image of a hard disk defects: (a) image before refinement; (b) image after refinement.

scratch, while in the initial image the borders of this region are blurred. A point-like defect, marked by a frame and readily seen in Fig. 8(b), is hardly noticeable in the initial image.

6. Conclusions

The coefficient of hard x-rays reflection from a supersmooth surface falls rapidly with roughness at grazing angles close to a critical angle. This effect may be used to visualize the roughness spatial distribution for such supersmooth flat surfaces, as semiconductor wafers, computer hard disks, x-ray and laser mirror substrates.

We have shown experimentally the possibility of the visualization of roughness spatial distribution with contrast high enough to detect roughness variations of about several Angstroms with a resolution of 1mm.

The application of the x-ray scanning method for the investigation of real objects, for instance, computer hard disks, proved the possibility of detecting various types of surface defects, such as roughness variation, scratches, point-like defects and curvature.

Acknowledgments

The authors express their gratitude to K. A. Valiev and to R. M. Imamov for their contribution to this work. We are especially grateful to S. A. Saunin and A. B. Shubin from MDT-NT for providing us with highly efficient software for image presentation. This work was financially supported by the Russian state programs "Advanced technologies and devices for micro- and nanotechnology," grant No.231/78/1-1, and "Scientific instruments," grant No.4/1.

References

Born, M. and Wolf, E., Principles of Optics, (1968), Ch.1, Pergamon, New York.

Gill, P. E., Murray, W., and Wright, M. H., Practical optimization, (1981), Ch.4, Academic Press, London.

Isakovich, M. A., The scattering of waves from statistically rough surface, Soviet Journal of Experimental and Technical Physics, 23-3(9)(1952-9), 305-314.

Névot, L. and Croce, P., Caractérisation des surfaces par réflexion rasante de rayons X. Application à l'étude du polissage de quelques verres silicates, Revue de Physique Appliquee, 15-3(1980), 761-779.

Protopopov, V. V., Valiev, K. A. and Imamov, R. M., X-ray scanner for the visualization of the spatial distribution of nanometer scale roughness, Proceedings of SPIE, 3275 (1998), 65-72.

Author Profile

Vladimir V. Protopopov: He graduated from the Moscow Institute of Physics and Technology (with honors) in applied physics in 1976. Received the degree of a candidate of sciences in radiophysics in 1978 and the scientific rank of senior scientist in 1983. He is a member of SPIE. His current scientific interests include x-ray optics, including focusing and collimating of beams, multilayer mirrors, microscopy. He is now with the Institute of Physics and Technology, Russian Academy of Sciences.

Vladimir V. Danilov: He graduated from the Moscow Power Engineering Institute in 1988. He received the degree of a candidate of sciences in electrotechnical systems control. He is now with the Institute of Physics and Technology, Russian Academy of Sciences.