Scanning Microscopy

Volume 1992
Number 6 *Signal and Image Processing in Microscopy and Microanalysis*

Article 39

1992

Holography - Just Another Method of Image Processing?

Hannes Lichte
Universität Tübingen Auf der Morgenstelle, Germany

Scanning Microscopy Supplement 6, 1992 (pages 433-440)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

0892-953X/92$5.00+.00

HOLOGRAPHY - JUST ANOTHER METHOD OF IMAGE PROCESSING?

Hannes Lichte

Institut für Angewandte Physik, Universität Tübingen
Auf der Morgenstelle 12, D-72076 Tübingen, Germany

Telephone Number: +49 7071 292428 / FAX number: +49 7071 295400

Abstract

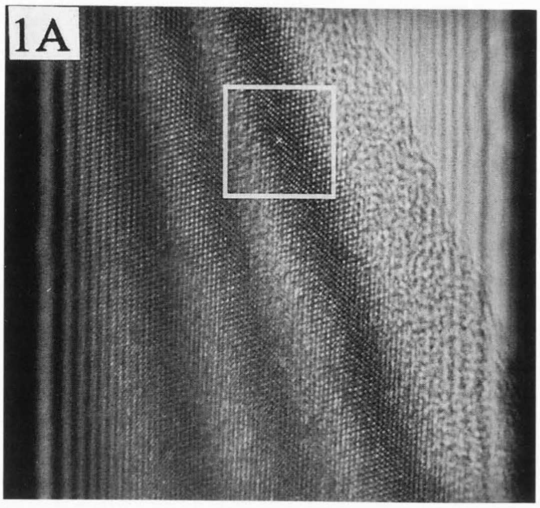
The advantage of all different kinds of electron holography lies in the fact that in addition to the image intensity they make use of the phase in real space and in Fourier space. Therefore, they are in principle superior to conventional imaging techniques, even if these are supported by highly sophisticated methods of image processing. The technique of electron off-axis holography, under development in Tübingen for more than 20 years, now enters atomic resolution domain. After correction of aberrations, amplitudes and phases can be determined quantitatively, both in real space and Fourier space, opening up novel facilities for the analysis of structures at atomic dimensions.

Key Words: Electron holography, electron wave optics, numerical wave reconstruction.

Introduction

In spite of the great success in all areas of natural sciences, transmission electron microscopy, from the viewpoint of optics, is still heavily restricted in its performance. Up to date, there are no satisfactory tools for the desirable analysis of the complex electron object wave by means of conventional microscopy. There is only one experimental parameter, i.e., focus, which can freely be selected to govern the transfer characteristics. However, even at Scherzer focus, the best choice to attain phase contrast at maximum point resolution, there is no large area phase contrast, resolution is severely limited, and amplitude and phase of the object wave show up in the image contrast indistinguishably mixed. Consequently, unique interpretation of a micrograph is generally very difficult.

Image processing techniques from an electron micrograph can only restrictedly help with these problems: it is true that, e.g., by means of a-posteriori Fourier filtering, even very weak crystalline structures can be detected more easily, and that noise reduction methods can be applied. However, one must keep in mind that, at the end, the processed image often represents less rather than more information about the object. More sophisticated methods, such as averaging over different samples of the same species, e.g., protein molecules, do indeed make a much better use of the information in the micrograph. To some extent, for weak phase objects, a correction for contrast reversal is possible, i.e., the regions with negative sign of the phase contrast transfer function $\sin\chi(R)$ can be rectified; however, because of the noise threshold, this holds only at some distance from the zeroes of the $\sin\chi$-function. Unfortunately, not even a simple measure like defocusing or correcting for some residual astigmatism can be performed. The basic reason is that the necessary information is not available from the micrograph, i.e., that the phase of the electron wave present in the final image plane cannot be recorded by conventional means. To overcome this limit, Gabor proposed electron holography: both the amplitude and phase of the electron wave are recorded in a hologram, from which it is reconstructed by light optical means using all the well-developed wave optical tools of light optics.



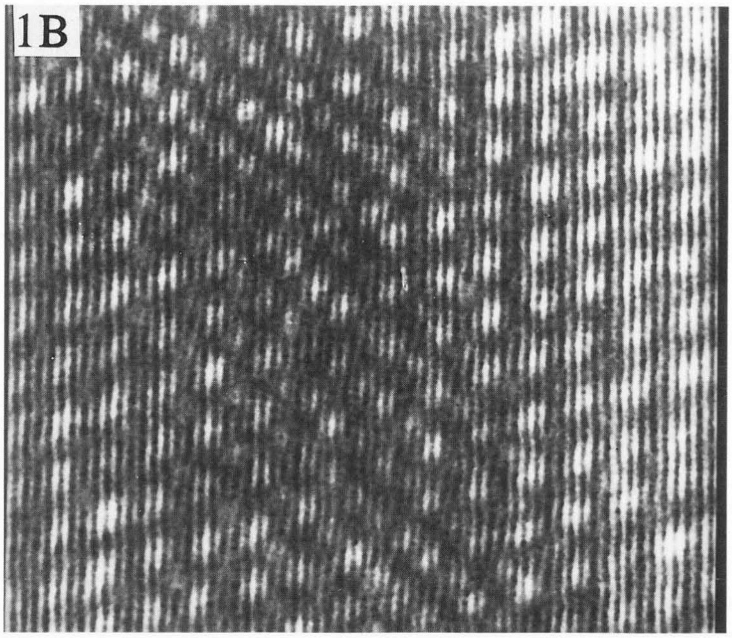


Figure 1. (**A**) Hologram of a wedge-shaped silicon crystal in (110) orientation. The thickness increases from top right to bottom left. Extinction contours show up. (**B**) At higher magnification, the hologram fringes (spacing 0.07 nm) show strong phase shifts, in particular at the extinction thickness.

We use the most powerful scheme of electron holography, called "image plane off-axis electron holography" as developed by Wahl [14], improved for the special needs of high resolution [3]. Instead of the proposed light optical reconstruction technique, we use numerical image processing methods. The hologram taken in the electron microscope is fed to a computer. Thereby, the computer is coupled wave-optically to the microscope; consequently, one is able to perform wave optical image processing for analysis of the electron wave which allows the use of all wave optical techniques one can think about. Since the two steps of taking the hologram and of subsequent evaluation are separated, they can be optimized individually.

Principles and Problems of High Resolution Holography

Taking the hologram

The image wave $b(x,y) = A(x,y) \cdot \exp(i\Phi(x,y))$ is coherently superimposed on a plane reference wave by means of the electron biprism [10] which yields an interference pattern called the hologram given by the intensity distribution

$$I_{hol}(x,y) = 1 + A^2(x,y) + 2A(x,y) \cdot \cos(2\pi R_c x + \Phi(x,y)) \quad (1)$$

An example is shown in Figure 1. In addition to the intensity $A^2(x,y)$ representing the conventional micrograph, one finds the cosinoidal term which contains the complete image wave imposed on a carrier wave with spatial frequency R_c. It is damped by the contrast V $(0 < V < 1)$ which has to be as good as possible to record the wave well above noise. V is given primarily by the degree of lateral coherence μ of the electron beam, by the instabilities and stray fields at the microscope, which can be described by a factor V_{inst}, and by the inelastic interaction that may occur as the electrons traverse the specimen. Electrons which have suffered an energy loss larger than 10^{-15} eV can no more be considered coherent with the reference wave, hence only elastic scattering information is collected in the hologram [4].

The contrast is given by

$$V = \mu \cdot V_{inst} \sqrt{1 - W} \quad (2)$$

with the probability (1-W) that an electron does not experience inelastic interaction [2].

The following items have to be taken into account to collect the maximum amount of information in the hologram:

Information limit R_{lim} of the image wave. The image wave $b(x,y) = A(x,y) \cdot \exp(i\Phi(x,y))$ is distorted with respect to the object wave $o(x,y)$ by convolution with the point spread function $FT\{E(R) \cdot \exp(i\chi(R))\}$ delivered by the objective lens

$$b(x,y) = o(x,y) \otimes FT\{E(R) \cdot \exp(i\chi(R))\} \quad (3)$$

The coherent aberrations (e.g., spherical aberration, defocus, astigmatism) described by the wave aberration (R) will be corrected numerically during reconstruction. However, this is only possible for spatial frequencies smaller than the information limit R_{lim}, i.e., as long as the wave is not damped below noise due to the incoherent aberrations, chromatic aberration and the illumination aperture; these are described by the corresponding damping function $E(R) = E_{chrom}(R) \cdot E_{ill}(R)$. Spatial frequencies higher than R_{lim} are not found in the image wave and they cannot be recovered by any means. Hence, the microscope has to be operated such that R_{lim} is sufficiently high. With a field emission electron gun, as a rule of thumb, R_{lim} is approximately twice as high as the point resolution at Scherzer focus.

Fringe spacing $1/R_c$. The higher the spatial frequency $R_{max} \le R_{lim}$ of the image wave one wants to reconstruct subsequently, the higher has to be the carrier spatial frequency R_c. Generally, as Wahl pointed out [14], the requirement $R_c \ge 3R_{max}$ has to be met. In any case, it does not make much sense to take holograms with a carrier frequency higher than $3R_{lim}$; one would not gain additional information but, instead, worsen the fringe contrast V, and hence damp the reconstructed wave, because the sensitivity to residual instabilities is heavily increased. In the case of our Philips 420/ST electron microscope, with a field emission gun, $R_{lim} \approx$ 6/nm, a fringe spacing of 0.05 nm is sufficient, 0.03 nm have been achieved [13].

Number of fringes. A certain minimum number of hologram fringes is needed for subsequent correction of aberrations. The reason is the following: By the coherent aberrations, the information stemming from an object point is smeared out in the image plane according to the point-spread-function $FT\{E(R) \cdot \exp(i\chi(R))\}$. For every object point to be reconstructed correctly, the point spread function has to be caught completely in the hologram. Consequently, the hologram must be wider than the width of the point spread function.

A closer look shows that the minimum width of the hologram must be broader than the point spread function by a factor of 4; otherwise, after Fourier transform, the wave aberration $\chi(R)$ is only defined with insufficient accuracy, i.e., worse than the Rayleigh limit $\pi/4$. The width w of the point spread function is, in geometrical approximation, related to $\chi(R)$ by $w = \text{grad}\chi(R)_{max}/\pi$, where $\text{grad}\chi(R)_{max}$ means the steepest gradient of $\chi(R)$ with respect to R in the range [0,Rmax], with $R_{max} \le R_{lim}$. Consequently, for an intended resolution R_{max}, the hologram must be made up by a number

$$N_{hol} \ge 4R_c w = 12 \cdot R_{max} \cdot w \qquad (4)$$

of fringes. Experimentally, the width of the hologram is limited by the degree of lateral coherence μ which can be realized. In principle, this does not present an obstacle, because a high degree of coherence can be set by means of the condenser lens. However, this is only possible at the cost of current density in the final image plane, because the coherent current $I_{coh}(\mu)$ available for taking a hologram at a degree of coherence μ

$$I_{coh}(\mu) = -B \cdot \ln(\mu)/k^2 \qquad (5)$$

(k = wave number); is limited by the brightness B of the gun. For very broad holograms, the small current density can be counterbalanced by long exposure times only restrictedly, because instabilities would again worsen the contrast of the hologram fringes.

Optimum focus. The only free parameter for taking a micrograph or a hologram is the focus. In conventional microscopy, a specific defocus is selected to optimize, e.g., the phase contrast. This is not important with holography, because phase and amplitude contrast are easily generated by the numerical processing later on. Therefore, the focus should be selected such that the hologram collects the maximum amount of information about the object. Points of concern are the information limit R_{lim} and the point spread function.

The envelope function of the illumination aperture

$$E_{ill}(R) = \exp(-\text{const} \cdot \theta_c^2(\text{grad}\chi(R))^2) \qquad (6)$$

depends on $\text{grad}\chi(R)$ with respect to R; θ_c means the semi-angle of the illumination aperture.

The width of the point spread function above was given as

$$w = (1/\pi) \cdot \{\text{grad}\chi\}_{max} \qquad (7)$$

where $\{\text{grad}\chi\}_{max}$ means the maximum absolute value of $\text{grad}\chi(R)$ in the spatial frequency range $[0,R_{max}]$.

For both aspects, the $\text{grad}\chi(R)$ plays the most important role. Therefore, recording the hologram can be optimized by operating the microscope at an optimum focus [5].

$$Dz = 0.75 \cdot C_s \cdot (R_{max}/k)^2 \qquad (8)$$

which minimizes the $\text{grad}\chi(R)$ function over the range $[0,R_{max}]$; C_s means the coefficient of spherical aberration of the objective lens, and k is the wave number of the electrons.

Numerical Reconstruction of the Hologram

Usually, the hologram is recorded on a photographic plate and, after digitization at 4096*4096 pixels by means of a charge-coupled-device (CCD)-line-camera, fed into the computer. Fourier transform yields, besides the diffractogram of the conventional micrograph $FT\{A^2(x,y)\}$ and a $\delta(R)$-peak on the optic axis, two redundant side-bands centered around $\pm R_c$:

$$\begin{aligned} FT\{I_{hol}(x,y)\} = \; & \delta(R) + FT\{A^2(x,y)\} + \\ & V \cdot FT\{A(x,y) \cdot \exp(i\Phi(x,y))\} \otimes \delta(R-R_c) + \\ & V \cdot FT\{A(x,y) \cdot \exp(-i\Phi(x,y))\} \otimes \delta(R+R_c) \end{aligned} \qquad (9)$$

One side-band is isolated. It represents the complex Fourier-spectrum of the image wave which compares well with the electron diffraction pattern found in the back focal plane of the electron microscope. The main advantage is that one has access to both amplitudes and phases, hence it represents an ideal starting point for wave optical image processing (Figure 2).

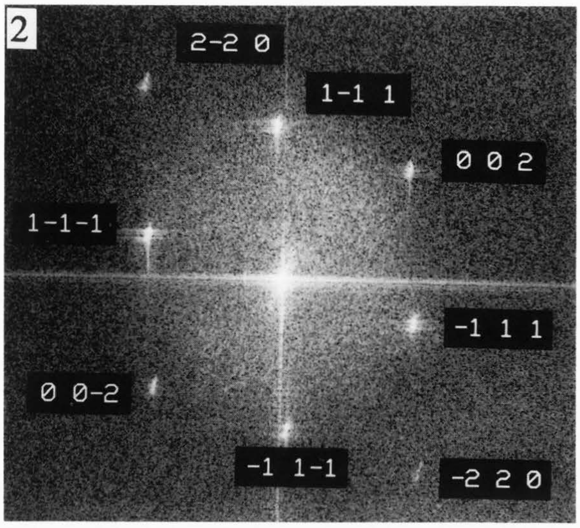


Figure 2. Fourier spectrum of the silicon crystal reconstructed from the hologram shown in Figure 1. The asymmetry in corresponding reflections, e.g., (2,-2,0) and (-2,2,0), suggests a slight mistilt of the crystal from the [110] zone axis.

Correction of aberrations

The Fourier spectrum represents the aberrated object wave, i.e., FT{o(x,y)} · exp(iχ(R)). It has to be divided by a phase plate exp(iχ(R)), to obtain, after inverse Fourier transform, the desirable object wave o(x,y) (Figure 3). By this technique, the coherent aberrations astigmatism, spherical aberration, and defocus have already been successfully corrected (Figure 4) [1]. Presently, we are working on the correction of the axial coma ("beam tilt"). The correction of the incoherent aberrations, i.e., compensating for the damping of the envelope function E(R), can be performed out to the information limit R_{lim}. Consequently, R_{lim} now represents the point resolution of the whole procedure.

Accuracy of correction

The aim of correction is to annihilate the wave aberration present in the microscope. Therefore, an appropriate phase plate has to be generated numerically given by:

$$\chi R = 2\pi k\,(0.25\,C_s\,(R/k)^4 - 0.5\,Dz\,(R/k)^2) \quad (10)$$

This phase plate has to agree with the wave aberration of the microscope for the range of spatial frequencies [0,R_{max}] with an accuracy better than $\pi/6$ to avoid residual blurring of amplitude and phase. This is by no means trivial, since, for the example of a 300 kV electron microscope with a spherical aberration of C_s = 1.2 mm and at optimum focus for R_{max} = 10/nm, the maximum range of the χ(R) function is approximately 24π. Therefore, the relative accuracy must be better than 0.007. Great care is required to establish the wave number k and the spatial frequency R with an accuracy

Figure 3. Numerically generated phase plates for correction of the wave aberration (Cs = 1.2 mm; Ub = 100 kV; bar = 2.5/nm). (**A**) For a hologram taken at defocus Dz = 0, strong moiré patterns arise because the number of pixels is too small to sample the steeply increasing wave aberration correctly. (**B**) At the optimum focus, in this case selected for R_{max} = 6.66/nm, the useful part of the phase plate reaches much higher spatial frequencies.

Figure 4. Holographic correction of aberrations without correction measures, both amplitude (**A**) and phase (**B**) show strong influence of astigmatism, spherical aberration and defocus. After correction, nearly all contrast is removed from the amplitude (**C**) whereas the object structure is mainly found in the phase image (**D**). This is to be expected since carbon foil is a weak phase object. Note the strong large area phase contrast in the phases indicating considerable thickness variations of the carbon foil; they do not show up in a conventional micrograph.

---

of a few tenths of a percent. Even if this is done perfectly, the spherical aberration and defocus Dz must be determined better than 4 μm and 0.5 nm, respectively [6]. This is not yet routinely possible.

Extraction of specific object information

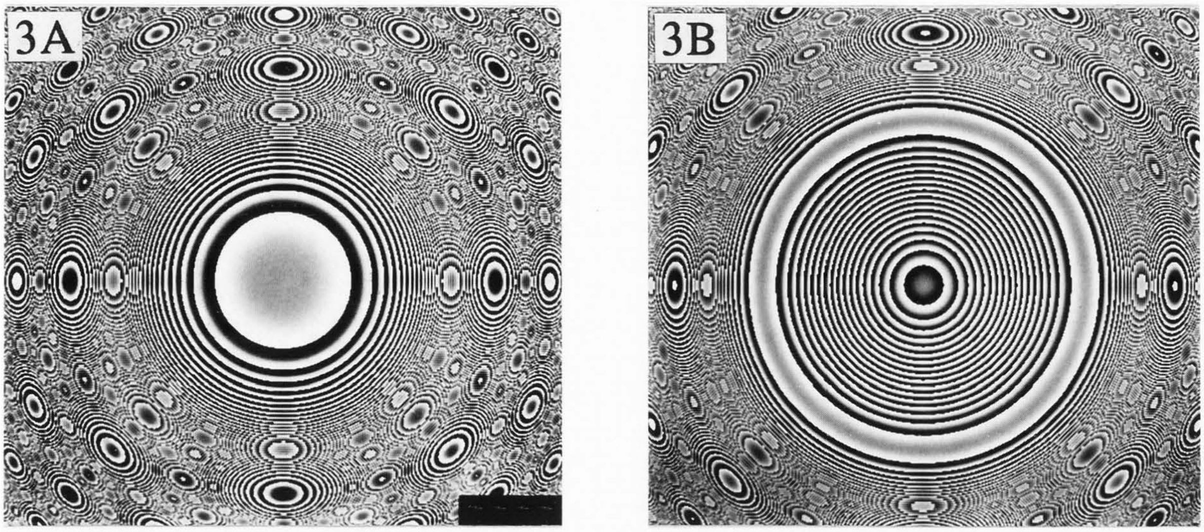
Since the Fourier spectrum is quantitatively available, there is a number of special methods at hand which allow the desired information about the object to be extracted. The most obvious one is to display, after inverse Fourier transform, amplitude and phase of the object wave. It can be performed under arbitrary focus (Figure 5) by applying a defocusing phase plate prior to the inverse Fourier transform.

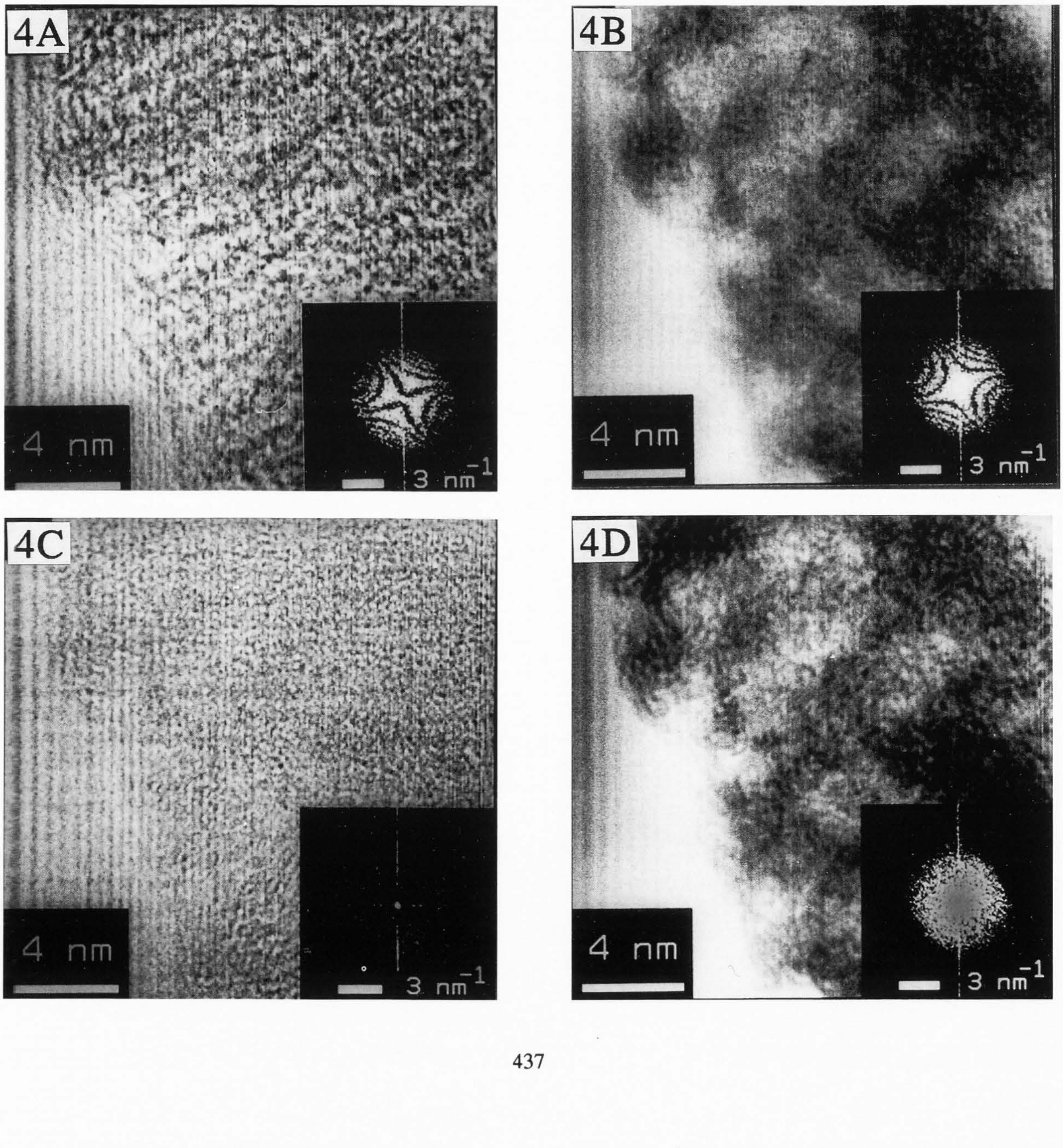
Analysis of amplitude and phase of the electron wave

Figure 6 shows amplitude and phase of an area of the silicon crystal which covers an extinction thickness. At the extinction thickness, due to the damping of the zero beam, half spacings show up as a dynamical effect. Fortunately, with holography there is no question whether or not these half spacings represent real lattice structures: they do not show up in the reconstructed phase, hence must be considered an thickness effect. This was to be expected from corresponding simulation calculations by means of the EMS program [12] for the wedge shaped silicon crystal [9].

Fourier-filtering

With the wedge shaped Si-crystal, the dependence of the complex scattering amplitudes on the thickness of the crystal can be measured in the following way: In the Fourier spectrum, single reflections are masked out by an aperture such that after inverse Fourier transform only the selected one contributes to amplitude and phase in real space; for example, this allows us to determine the variation of amplitude and phase of the scattering waves with the thickness of a crystal (Figure 7) [8].


3A
3B


4A
4 nm
3 nm^{-1}
4B
4 nm
3 nm^{-1}
4C
4 nm
3 nm^{-1}
4D
4 nm
3 nm^{-1}

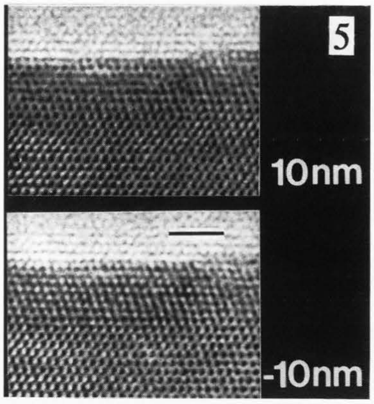
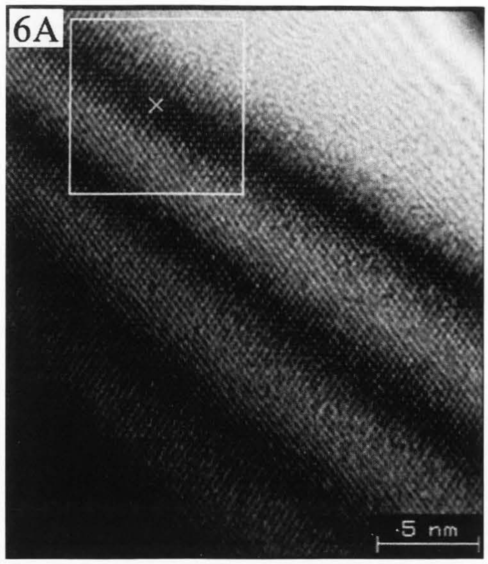


Figure 5. Effect of a posteriori focusing of the electron wave reconstructed from a hologram. Changing the focus by means of application of a corresponding phase plate prior to the inverse Fourier transform produces the same phenomena as in the microscope, e.g., contrast changes [7]. Bar = 3 nm.



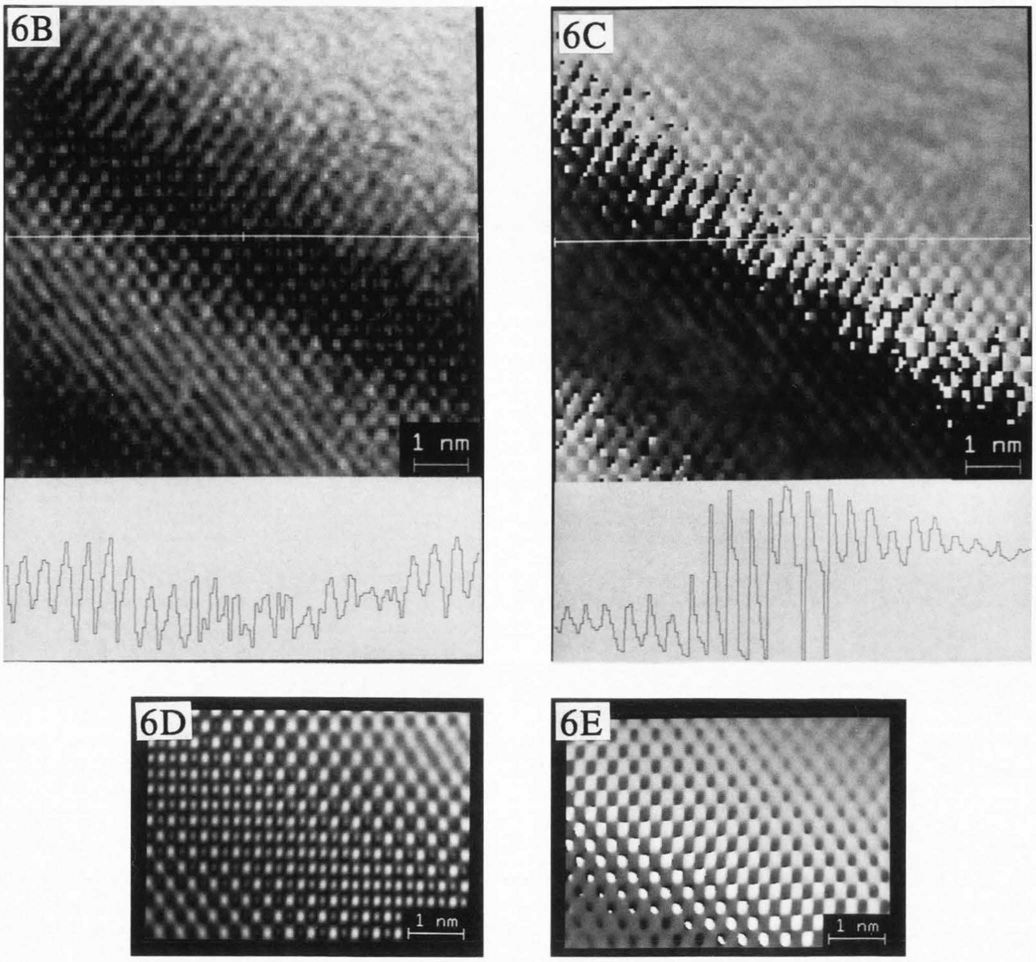


Figure 6. (**A**) Electron wave transmitted through a wedge shaped silicon crystal. The thickness increases from top right to bottom left. (**B**, **C**): At the extinction thickness, half spacings show up in the amplitude (**B**), however, not in the phase image (**C**). As displayed by the line scan, the phase modulation due to the Si atoms steadily increases on either side of the extinction thickness, where a sudden phase jump occurs (full scale 2π). (**D**, **E**): The corresponding results simulated with EMS compare well with the reconstructed wave. From reference [9].

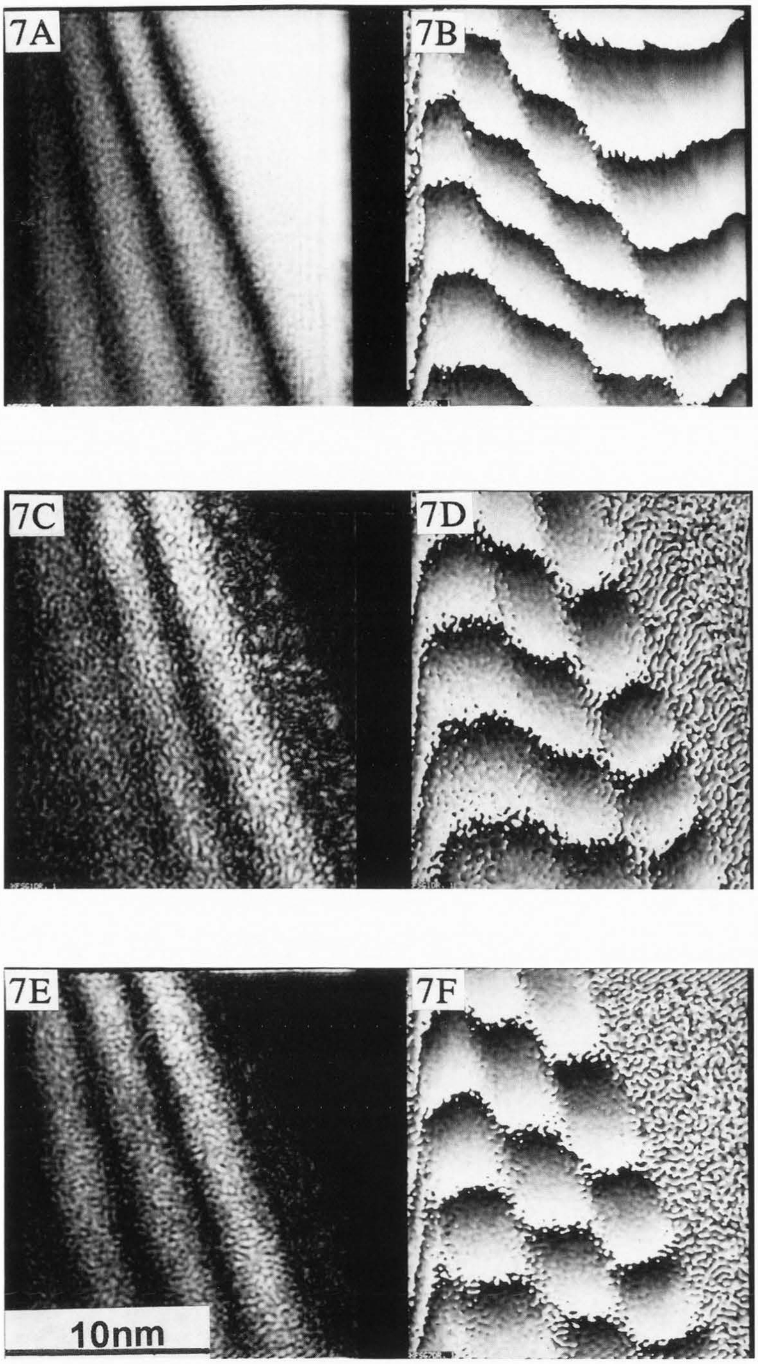


Figure 7. Amplitude (A-C) and phase (D-F) of the waves reconstructed from single reflections: A: (0,0,0); B: (1,-1,1); and C: (1,-1,-1). At the extinction thicknesses, contour lines and phase jumps show up in amplitude and phase, respectively. By comparison with computer simulation, the center of the Laue circle was estimated at (0.075,-0.075,-0.227) instead of (0,0,0) [9].

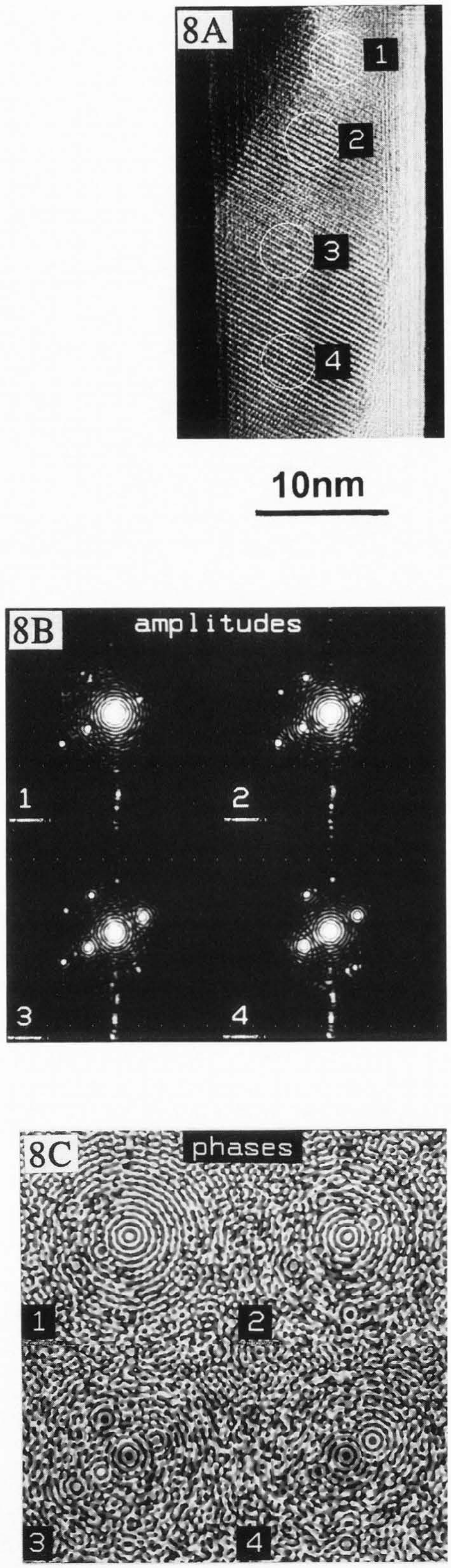


Figure 8. Holographic nanodiffraction. (**A**) A numerical selected-area-aperture is applied at different positions of the niobium-oxide crystal. After inverse Fourier transform, the diffraction patterns can be analyzed both in amplitude (**B**) and phase (**C**). The strong change in symmetry and intensity of the reflections from position 1 to 4 indicates some twist of the crystal. The aperture diameter is 4 nm; apertures with diameter less than 1 nm have already been tested.

Additionally, in a very flexible way, apertures of arbitrary size and shape can be applied to the spectrum such that any kind of investigation can be performed, e.g., by means of dark field imaging or lattice imaging with an arbitrary selection of reflections.

Nanodiffraction

If in conventional microscopy one is interested in diffraction analysis of small selected areas of the object, one has to apply the selected area aperture and switch to diffraction mode. Unfortunately, this is not possible at high resolution, because the smallest aperture available, referred back to the object, is approximately 100 nm in diameter. If one tries to perform this kind of analysis from a micrograph, e.g., by laser diffractometry, then the diffractogram obtained does not present the information one is interested in: for example, a diffractogram is always symmetric with respect to the optic axis, hence crystal tilt cannot be examined from symmetry; furthermore, a diffractogram only reveals the structures which are converted into a contrast in the micrograph. Figure 8 shows that with holography these restrictions are overcome. The complex diffraction pattern from areas only a few nanometers wide can be analyzed.

Summary

Electron holography opens up new dimensions in image processing: Via the hologram, the electron wave is transferred to the computer enabling coherent processing of the electron wave. As an extension of the microscope, the computer allows the electron wave to be analyzed without the heavy technical restrictions one faces when operating the microscope. Some of these processing steps can already be performed in real-time [11].

Since for a thorough analysis of an object all the data can be retrieved from a single hologram, they represent the very same object in that absolutely no change occurs during the analysis, neither by instabilities of the microscope nor by alterations of the specimen, say, by radiation damage.

Acknowledgements

This paper broadly summarizes results achieved so far by the joint efforts of the electron holography group in Tübingen. The financial support obtained from Körber-Stiftung, Volkswagen-Stiftung, Deutsche Forschungsgemeinschaft and the European Community is gratefully acknowledged.

References

[1] Fu Q, Lichte H, Völkl E (1991) Correction of aberrations of an electron microscope by means of electron holography. Phys. Rev. Lett. **67**, 2319-2322.

[2] Harscher A, Lenz F, Lichte H (1992) Electron holography provides zero-loss-images. Proc. EUREM'92, Granada, Last Minute Brochure, p. 35 (copy available from H. Lichte).

[3] Lichte H (1986) Electron holography approaching atomic resolution. Ultramicrosc. **20**, 293-304.

[4] Lichte H (1991) Electron Image Plane Off-Axis Electron Holography of Atomic Structures. Adv. in Opt. and Electr. Microsc. **12**, 25-91.

[5] Lichte H (1991) Optimum focus for taking electron holograms. Ultramicrosc. **38**, 13-22.

[6] Lichte H (1992) Can electron holography reach 0.1 nm resolution? Ultramicrosc. **47**, 223-230.

[7] Lichte H, Völkl E (1991) Electron holography in material science. Proc. EMSA, San Francisco Press, 670-671.

[8] Lichte H, Völkl E, Scheerschmidt K (1992) First steps of high resolution electron holography into materials science. Ultramicrosc. **47**, 231-240.

[9] Lichte H, Orchowski A, Scheerschmidt K (1992) Electron holography for analysis of a wedge-shaped silicon crystal. Proc. EUREM'92, Granada, Last Minute Brochure, p. 31 (copy available from H. Lichte).

[10] Möllenstedt G, Düker H (1956) Beobachtungen und Messungen an Biprisma-Interferenzen mit Elektronen-wellen (Observations and measurements of biprism interferences with electron waves). Z. Physik **145**, 377-397.

[11] Rau WD, Lichte H, Völkl E, Weierstall U (1991). Real-time reconstruction of electron off-axis holograms recorded with a high pixel CCD camera. J. Computer-Assisted Microscopy **3**, 51-63.

[12] Stadelmann P (1987) EMS - a software package for electron diffraction analysis and HREM image simulation in materials science. Ultramicrosc. **21**, 131-145.

[13] Völkl E, Lichte H (1990) Electron holograms for subangstrom point resolution. Ultramicrosc. **32**, 177-180.

[14] Wahl H (1975). Bildebenenholographie mit Elektronen (Image plane holography using electrons). Thesis, Univ. Tübingen, Germany.

Discussion with Reviewers

J.M. Rodenberg: Presumably the question of the number of fringes required in the hologram for satisfactory deconvolution can be thought of as a windowing problem in real space.
Author: Yes.

J.M. Rodenberg: In performing the unaberrated restoration, does the author use a smooth window function in real space to avoid truncation errors?
Author: For certain reconstruction steps, e.g., quantitative evaluation of the data in Fourier space, soft-windowing is indispensable because otherwise falsification due to streaking effects may occur.