A new approach to analysis of subwavelength sized secondary light sources

Nikolay B. Voznessenski, Vadim P. Veiko, Vitaly M. Domnenko, Alexey E. Goussev, Tatyana V. Ivanova and Sergey A. Rodionov

St.-Petersburg State Institute of Fine Mechanics and Optics (Technical University), Sablinskaya ul.,14, St.-Petersburg, Russia, 197101

ABSTRACT

A new approach to the investigation of probes for scanning near-field optical microscopes (SNOM) and recognition of parameters of arbitrary secondary light sources in nanometric scale is suggested. A new numerical technique of analytical continuation of the Fourier spectrum with the object restoration procedure based on Zernike polynomials iterative extrapolation is presented.

Keywords: near-field optics, subwavelength objects, Fourier transform, superresolution.

1. INTRODUCTION

Nowadays a great amount of tasks in physics, biology, medicine and technology deal with analyzing nanometer sized objects and structures namely photomasks and compact optoelectronic devices, nanopowder particles, nanosubstrates, tools for micro and laser surgery, biomolecules, viruses, etc. In all these cases the urgency of the use of optical research methods is doubtless. However it is evident that the far-field optical microscopy offers no direct characterization technique in nanometric scale and it is necessary to take alternate means such as electron scanning microscopy techniques. Some of tasks in question accommodate the use of the latter methods but there are cases in which their application is hardly reasonable and even impossible. This concerns objects which cannot be observed with the use of electron irradiation for example the objects imposing only visual electromagnetic field, being investigated in vivo, interesting only by their optical properties, etc. So the near-field optics with the variety of its opportunities becomes more and more preferable. It is important that the light scattering on extra small material structures may be used according to the Babinet's principal for investigating amplitude and phase effects imposed by these structures. Hence the possibility appears of the knowledge of their shape and sizes as the parameters of some secondary sources of light.

Here the following problems take place:

1) investigation of the light field distribution (amplitude and phase) quite close to the secondary source for which new extra sensitive holographic receivers or CCD matrices are necessary;

2) reconstruction of amplitude and phase distribution on the source by the knowledge of its far-field radiation.

The first problem is being solved by application of a set of scanning near-field optical microscopy (SNOM) techniques. This field as a fast growing scientific branch offers a number of new ideas which may be applied in other disciplines. This paper is devoted to the secondly mentioned problem and at the same time to the problem of estimation of SNOM scanning apertures by investigation of far-field light distribution. Some time ago Chr. Obermüller and Kh. Karrai[1] had investigated the possibilities of far-field registration in order to define the aperture of a secondary light source in submicron scale. This task leads to a generally formulated problem of investigating an arbitrary separate subwavelength sized optical object by its far-field diffracted light processing. In this case the shape recognition with great lack of information meets great troubles

---

Further author information -
N.B.V.: E-mail: vnb@aco.ifmo.ru
V.P.V.: E-mail: veiko@lastech.ifmo.ru

and requires a preliminary mathematical extension of the registered part of the diffracted field distribution in order to grasp nanometric dimensions and anything differing from a simple circle.

2. MATHEMATICAL MODEL OF THE NEAR AND FAR FIELDS

The theoretical and mathematical problems come from the difficulties of far- and near-field phenomena representation in one model in terms of linear reversible equations. The suggested mathematical modeling is based on the superposition of solutions of Maxwell's equations enclosing linearly polarized vector plane waves in real and complex forms as functions of spatial frequencies $\left(\nu_x, \nu_y\right)$ corresponding both to the propagating $\mathbf{u}_{0ij}^{R} = \dfrac{\mathbf{k}_{ij} \times \mathbf{p}_0}{\left|\mathbf{k}_{ij} \times \mathbf{p}_0\right|}$ and evanescent $\mathbf{u}_{0ij}^{E} = \dfrac{\hat{\mathbf{k}}_{ij} \times \mathbf{p}_0}{\left|\hat{\mathbf{k}}_{ij} \times \mathbf{p}_0\right|}$ light. Here $\mathbf{u}_{0ij}^{R}, \mathbf{u}_{0ij}^{E}$ are the complex vector amplitudes, real $\mathbf{k}_{ij}$ and complex $\hat{\mathbf{k}}_{ij}$ denote the wave vectors corresponding to the space frequencies below and higher than $\dfrac{1}{\lambda}$, $\mathbf{p}_0$ is the polarization status vector.

The expression for the entire near-field distribution close to the investigated light source will be:

$$\mathbf{U}(\mathbf{r}) = \sum_i \sum_j \mathbf{B}_{ij} \mathbf{u}_{ij}(\mathbf{r}),$$

where $\mathbf{B}_{ij}$ – matrices of complex coefficients, $\mathbf{u}_{ij}(\mathbf{r}) = \begin{cases} \mathbf{u}_{ij}^{R}(\mathbf{r}) \\ \mathbf{u}_{ij}^{E}(\mathbf{r}) \end{cases}$, $\mathbf{u}_{ij}^{R}$ – propagating plane waves, $\mathbf{u}_{ij}^{E}$ – evanescent plane waves.

The use of vector representation of the light complex amplitude leads to effective calculation procedures for rigorous far-field intensity distribution simulation. If the vector expressions for both types of plane waves are substituted into the above sum with taking into account that $\mathbf{k}_{ij}$ and $\hat{\mathbf{k}}_{ij}$ are the functions of spatial frequencies $\left(\nu_x, \nu_y\right)$ it will be possible to use the digital Fourier transform to calculate all the components of the sum. The description of the vector amplitude distribution itself may be defined in the following form: $\mathbf{U}(\mathbf{r}) = \mathbf{S} \cdot \mathbf{u}_0$, where $\mathbf{S}$ – is the matrix model of the light source containing its shape and sizes, $\mathbf{u}_0$ – vector amplitude of an incident linearly polarized plane wave with the wave vector $\mathbf{k}_{00}$. Using this matrix expression together with the sum for $\mathbf{U}(\mathbf{r})$ and vector amplitudes for plane waves the following matrix equation may be obtained:

$$\mathbf{S} = F\left[\mathbf{P}_{ij}\left(\nu_x, \nu_y\right)\mathbf{B}_{ij}\right],$$

where F – direct Fourier transform operator, $\mathbf{P}_{ij}$ – a set of matrix operators of rotations according to the assumed definition of a plane wave vector amplitude, $\mathbf{B}_{ij}$ – a set of matrix coefficients, indices ij denote that the Fourier transform is taken with respect to the digitized spatial frequencies.

The matrix expression for the far-field light amplitude distribution parameters is $\mathbf{B}_{ij} = \mathbf{P}_{ij}^{-1}\left(\nu_x, \nu_y\right) \cdot F^{-1}[\mathbf{S}]$, where $\mathbf{S}$ is the mentioned matrix of the secondary light source parameters, $\mathbf{P}_{ij}^{-1}\left(\nu_x, \nu_y\right)$ – a set of inverse matrix operators of rotations, F^{-1} – the inverse Fourier transform operator. Principal advantage of this model is the use of a set of linear reversible transformations based on the Fourier integral. If the amplitude of the far-field light distribution is known it will be possible to restore the initial light field distribution encoded in the matrix $\mathbf{S}$ by the inverse matrix formula for the restored initial distribution $\mathbf{S}'$.

The registered far-field radiation is the angular intensity distribution with respect to angles of diffraction $I'\left(\theta_x, \theta_y\right)$ and that is why it should be transformed into the complex amplitude. This transformation leads to the loss of phase which fortunately may be taken as a constant. But the knowledge of the incident light polarization must be taken into account. As it is often noted the polarization status remains without any change inside the fiber as well as while scattering on the aperture. Therefore the vector amplitude of the registered light may be described as follows:

$$\mathbf{U}'\left(\theta_x, \theta_y\right) = \begin{pmatrix} u'_x\left(\theta_x, \theta_y\right) \\ u'_y\left(\theta_x, \theta_y\right) \\ u'_z\left(\theta_x, \theta_y\right) \end{pmatrix} = \mathbf{B}' \cdot \frac{\mathbf{k}\left(\theta_x, \theta_y\right) \times \mathbf{p}_0}{\left|\mathbf{k}\left(\theta_x, \theta_y\right) \times \mathbf{p}_0\right|},$$

where $\mathbf{p}_0$ is the polarization status vector, $\mathbf{B}' = \begin{pmatrix} \sqrt{I'\left(\theta_x, \theta_y\right)} & 0 & 0 \\ 0 & \sqrt{I'\left(\theta_x, \theta_y\right)} & 0 \\ 0 & 0 & \sqrt{I'\left(\theta_x, \theta_y\right)} \end{pmatrix}$ is the diagonal matrix containing the square roots of the registered values of the angular intensity distribution.

This equation is correct in case of linear polarization and that the electric vector lies in the plane of a receiver. The latter may be provided if the receiver is being set at a proper angle while scanning or the field is registered in the focal plane of a microobjective with a high numerical aperture.

Now for the restored boundary conditions matrix $\mathbf{S}'$ after changing angular variables to spatial frequencies we evidently obtain the following expression:

$$\mathbf{S}' = F\left[\mathbf{P}_{ij} \cdot \mathbf{B}'\right].$$

where F – direct Fourier transform, $\mathbf{P}_{ij}$ – a set of matrices of rotations.

The restored boundary conditions $\mathbf{S}'$ may differ greatly from the initial ones $\mathbf{S}$ because $\mathbf{B}'$ does not contain direct information about the near-field distribution. The difference increases as a subwavelength sized hole diminishes further. This comes from that the share of propagating waves in the whole mathematical spectrum extremely diminishes because the spatial frequencies of the propagating waves occupy a small interval. To improve the task of restoration of the secondary light source parameters an additional mathematical procedure is required.

3. APPROACHES TO ANALYTICAL CONTINUATION OF FOURIER SPECTRA

The task of more detailed definition of the input distribution may be solved with the use of its continuated Fourier spectrum. The process of continuation itself usually is being built as the process of recognition of the input with a sequence of feedback procedures. The real examples of such kind deal either with the analytical continuation of the spectrum on the basis of the sample theorem[2] or with various iterative procedures[3]. Also there is a theoretical approach with the use of wave prolate functions with double orthogonality[4] being applied to antenna currents distributions. In case of two dimensions which is in optics this approach meets with the problem of two-dimensional basis construction and a great amount of calculations.

Both approaches - the sample theorem and iterative procedures give satisfactory results with greater or less probability but in our case it is necessary to obtain the stable information about shape parameters. The better stableness may come from the use of functions with double orthogonality where the expansion of the spectrum by these functions may be immediately defined by quite correct approximation procedures. Since such expansions have been successfully used for one dimensional space they were not in use in optics. In this work a new analytical-iterative procedure is offered. This procedure involves a step of polynomial expansion which is realized as Zernike polynomial extrapolation with new approach to calculation of polynomials beyond their orthogonality region with high accuracy. This step gives a set of orthogonal polynomial expansion coefficients which form a numerical model of the visible part of the spectrum strongly related with the model of the invisible one. With these coefficients and if the analytical expressions of the functions with double orthogonality with two dimensions were known and suitable for computing the task of restoration of subwavelength sized shapes would be solved at once. Unfortunately there are no computable analytical expressions to continue the spectrum in the invisible region by these functions. That is why it is reasonable to carry out a sequence of approximations in order to find the function mostly close to the part of the spectrum already defined. In our approach the use of sample functions in an iterative process has the sense of orthogonal expansion of the spectrum and should be performed more surely. The calculations taken on Pentium-166 has shown that this idea was quite correct.

As above mentioned the wave prolate functions which are also known as the spheroidal functions with double orthogonality may be expressed through orthogonal polynomials within the visible region and through infinite functions with finite spectra beyond this region. In case of one dimension the Legendre polynomials are used for the visible part and the Bessel functions of the integer order - for the invisible one. Besides each Legendre polynomial must be connected to its own order Bessel function. The similar situation is with two and more dimensions only the polynomials for the visible part must be generally orthogonal and the functions for the invisible part are not Bessel functions but the solutions of the general integral equation:

$$\psi(\mathbf{x})\lambda = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \psi(\mathbf{x}')\exp\left[2\pi i\left(\mathbf{x}'^{\mathrm{T}}\mathbf{x}\right)\right]d\mathbf{x}' ,$$

where $\mathbf{x}$ and $\mathbf{x}'$ are the vectors in the n-dimensional space.

The mathematical analysis[4] shows that these functions compile an orthogonalized set of functions with finite spectra and region of orthogonality covering the whole invisible region which ends at infinity. Instead of deriving analytical expressions for these functions promising to be much more complicated and hardly computable than in one-dimensional case we may construct these functions with a simple kind of orthogonal ones, for example, sample functions. That is why any iterative procedure of superresolution may be taken for quite correct numerical continuation of the invisible part of the spectrum.

After processing the registered intensity $I'(\nu_x, \nu_y)$ the visible part of the Fourier spectrum of the initial conditions $f^{v}(\nu_x, \nu_y) = Bf(\nu_x, \nu_y)$ is to be defined. Here B denotes that the total spectrum $f(\nu_x, \nu_y)$ is being cut off by the circle $\left(-\frac{1}{\lambda} \le \sqrt{\left(\nu_x^{\,2} + \nu_y^{\,2}\right)} \le \frac{1}{\lambda}\right)$. The mathematical equations showing the process of continuation may be divided into two groups according to the steps of continuation.

THE FIRST STEP (ZERNIKE EXTRAPOLATION)

The formula for polynomial extrapolation of the spectrum outside the visible region is

$$f^{v+}(\nu_x^{\,+}, \nu_y^{\,+}) = \sum_i \sum_j P_{ij}(\nu_x^{\,+}, \nu_y^{\,+})C_{ij} ,$$

where C_{ij} are the Zernike polynomial expansion coefficients satisfying the equation $f^{v}(\nu_x, \nu_y) = \sum_i \sum_j P_{ij}(\nu_x, \nu_y)C_{ij}$. Here $P_{ij}(\nu_x, \nu_y)$ are the orthogonal Zernike polynomials on the region $\left(-\frac{1}{\lambda} \le \sqrt{\left(\nu_x^{\,2} + \nu_y^{\,2}\right)} \le \frac{1}{\lambda}\right)$ and $\left(\nu_x^{\,+}, \nu_y^{\,+}\right)$ denote the coordinates of points in the region greater than the initial visible space. The cross dimensions of this new region may be 1.5, 2. and more times larger than the old one. The greater the relative dimensions of the new region are the iller-posed the task is. In order to improve the solution of this task a new approach to the polynomial extrapolation is suggested. The polynomials should be built as special orthogonal polynomials with variable properties of orthogonality. Such polynomials had been elaborated in the St.-Petersburg Institute of Fine Mechanics and Optics in 1983 in order to enrich the properties of Zernike polynomials by new numerical advantages[5].

The main idea of extrapolation is quite simple and is based on follows: 1)the polynomials should never be calculated beyond the unit circle which can be maintained by scaling, 2)the diameter of orthogonality region may be 0.8, 0.5 and even smaller than the region of extrapolation. In this case the error of polynomial expansion (of the visible part of the spectrum) within the orthogonality region should be very small and the following extrapolation error within the unit circle be rather small too. Such an idea offers a sufficiently long extrapolation. The expansion coefficients C_{ij} may be defined in the visible region by the very stable standard Gram-Schmidt procedure.

THE SECOND STEP (SUPERRESOLUTION PROCEDURE)

Since the result of the first step is the extended "visible" region the superresolution procedure can be run under promoted conditions. Consequently one of the well-known iterative approaches[3,7] may be taken as a means to complete the continuation. In our case this process is formulated with establishing the feedback in the spectrum space as follows:

$$s_e^{(p)}(x,y) = F\left[f^{v}\left(\nu_x,\nu_y\right)\right],$$

where $s_e^{(p)}(x,y)$ is the preliminary estimation of the input. This estimation is being analyzed in order to find the smallest region Ω_s enclosing all the points where the values of the signal exceed a taken numerical threshold s_0. The signal $s_e^{(p)}(x,y)$ within Ω_s is denoted as $Qs_e^{(p)}$ and the first estimation of the spectrum will be

$$f_e\left(\nu_x,\nu_y\right) = F^{-1}\left[Qs_e^{(p)}(x,y)\right].$$

This is a description of the spectrum by a set of two-dimensional sample functions with proper configuration obtained by means of digital Fourier transform.

Then after the invisible part of $f_e\left(\nu_x,\nu_y\right)$ is added to the given "visible" spectrum the following transformation is to be performed:

$$s_e^{(s)}(x,y) = F\left[f_e\left(\nu_x^{+},\nu_y^{+}\right) + f^{v}\left(\nu_x^{-},\nu_y^{-}\right)\right],$$

where $s_e^{(s)}(x,y)$ is the secondary estimation of the input, $\left(\nu_x^{-},\nu_y^{-}\right)$ define the points of the "visible" part of the spectrum and $\left(\nu_x^{+},\nu_y^{+}\right)$ define the points of the rest area. Then $s_e^{(s)}(x,y)$ should be substituted instead of $s_e^{(p)}(x,y)$ into the equation for $f_e\left(\nu_x,\nu_y\right)$ and the iterative cycle be continued.

The process is going on until the value of

$$\left(\iint_{\Omega_e}\left|f_m\left(\nu_x^{-},\nu_y^{-}\right) - f^{v+}\left(\nu_x^{-},\nu_y^{-}\right)\right|^2 d\omega\right) \Big/ \iint_{\Omega_e}\left|f^{v+}\left(\nu_x^{-},\nu_y^{-}\right)\right|^2 d\omega$$

becomes less than a chosen tolerance ε which usually depends on the numerical threshold s_0, $f_m\left(\nu_x^-, \nu_y^-\right)$ being the spectrum of the m-th estimation of the input calculated within the "visible" region.

This approach has been tested for various shapes of holes with different dimensions for various sets of orthogonal polynomials in order to optimize the process. The numerical calculations have shown that this technique may give very promising results of sufficiently long continuation of the spectrum.

4. THE RESULTS OF SHAPE PARAMETERS RESTORATION FOR ELLIPTICAL AND RECTANGULAR APERTURES

The suggested concept of relation between near- and far-field distributions has been applied to a binary model of a secondary light source with circular, elliptical and rectangular apertures. It is the model of diffraction of a normally incident linearly polarized plane wave on a screen with a subwavelength sized hole. This example offers to characterize in detail the principal opportunities and accuracy of a suggested restoration technique. Firstly a far-field distribution of light scattered on a hole with taken shape and dimensions was calculated with taking into account only frequencies of the propagating waves. Secondly after this distribution had been interpreted as a very small part of the Fourier spectrum according to the described mathematical transformations the restoration procedure was run. The taken apertures offer to show not only the opportunities of size recognition but the profile as well. During calculations it was found that simple smallest shapes may be restored with exclusively high accuracy which may help to organize a very stable process of SNOM optical probes estimation.

The results of shapes restoration are presented for a rectangular aperture set in the (x, y) plane at an angle of 45° and the possibly least object - an aperture formed by 8 pixels. The investigation of the shape estimation error had been carried on by several methods involving different levels of iteration. The difference between them depends on a taken set of polynomials and length of extrapolation. The longer extrapolation makes it easier to run the superresolution procedure but unfortunately reduces the correctness. Figures 2-5 and 7-10 display the numerical model of near-field intensity distributions restored for certain secondary sources, all of these distributions having the boundary profile closer to the initial shape of a source aperture while coming to a higher level of iteration. The minimum of error is determined by the sizes of a pixel. Obviously in order to diminish it a larger sample should be taken. The table below shows the reducing of a profile error corresponding to a certain level of iteration and therefore it may be derived that any enhancement of the iteration procedure leads to higher resolution. All the calculations have been performed for a wavelength of 500 nm and the angle of observation $\pm 71^\circ$, the sample dimensions were taken 512×512. In order to diminish the minimum of the investigated aperture the sample 1024×1024 or 2048×2048 should be taken which requires the additional memory of 64 MB on PC.

5. CONCLUSION

The results show that the restoration of the initial shape of a source aperture may be successfully obtained only with the use of preliminary polynomial extrapolation. The latter offers a reasonable promotion of the iterative procedure with a considerable regress of profile error. So the presented approach demonstrates promising possibilities of mathematical processing of image reconstruction in case of too long continuation of its spectrum. This method may be used firstly in near-field scanning microscopy for probes apertures and SNOM transfer function estimation. However the task of a source aperture parameters estimation by the far-field characteristics cannot be said to be entirely solved because of two main problems: 1) a problem of the conditioning of this task which may be checked by further numerical investigations with the use of noised "measured" intensity data; 2) a problem of $<0.05\lambda$ sized apertures analysis. The latter requires not only the additional memory but a higher level of iteration as well. The further work will be devoted to the detailed investigation of the degree of conditioning and other levels of iteration and it is worth to say that the suggested approach is promising to offer a satisfactory recognition of arbitrary small apertures.

6. AKNOWLEDGEMENTS

The authors wish to thank G. Sepold, W. Jüptner, W. Osten and S. Seebacher (Bremen Institute of Applied Beam Technology) for useful and illuminating discussions.

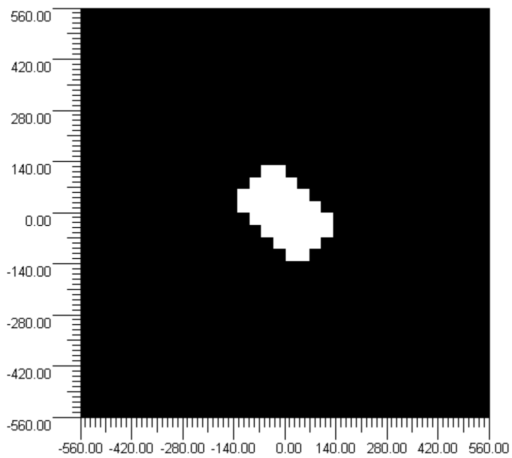

Fig.1 The rectangular aperture 280×180 (nm) at an angle of 45°.

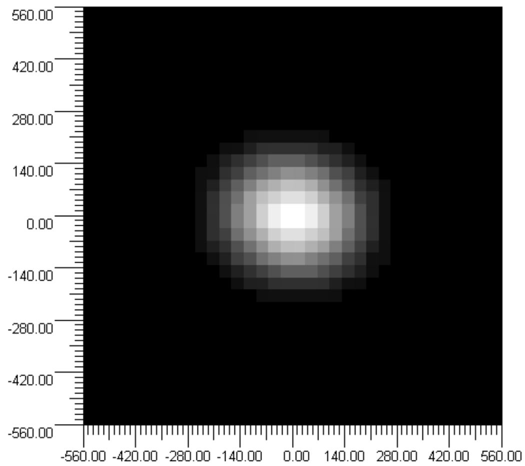

Fig.2 The near-field intensity distribution restored by the 0-iteration.

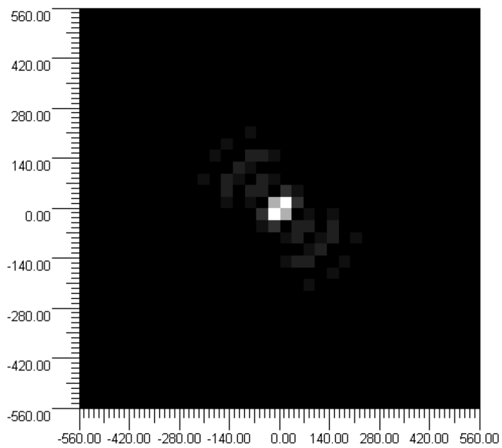

Fig.3 The near-field intensity distribution restored by the 1st iteration.

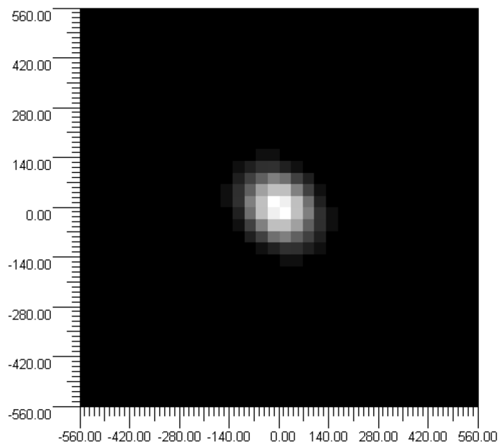

Fig.4 The near-field distribution restored by the 2nd iteration.

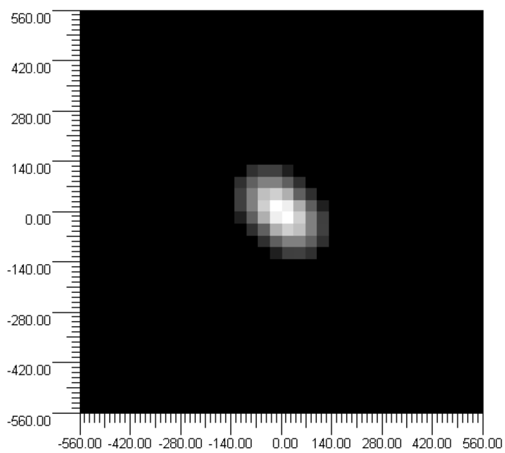

Fig.5 The near-field intensity distribution restored by the 3rd iteration.

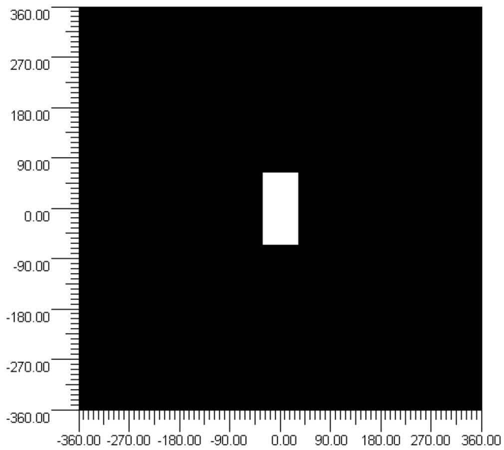

Fig.6 The rectangular aperture 90×180 (nm).

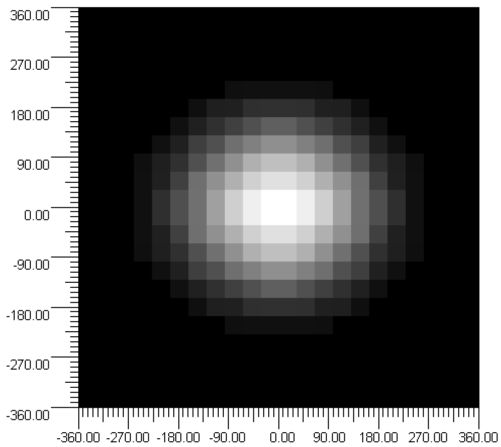

Fig.7 The near-field intensity distribution restored by the 0-iteration.

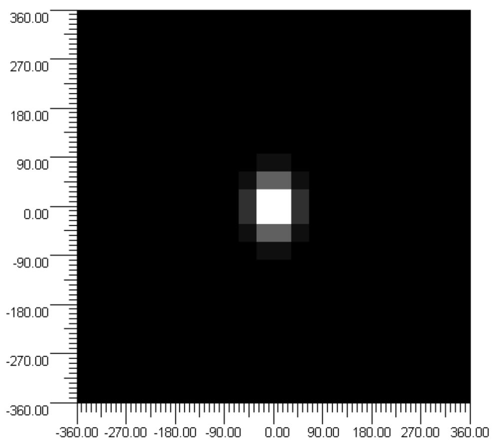

Fig.8 The near-field intensity distribution restored by the 1st iteration.

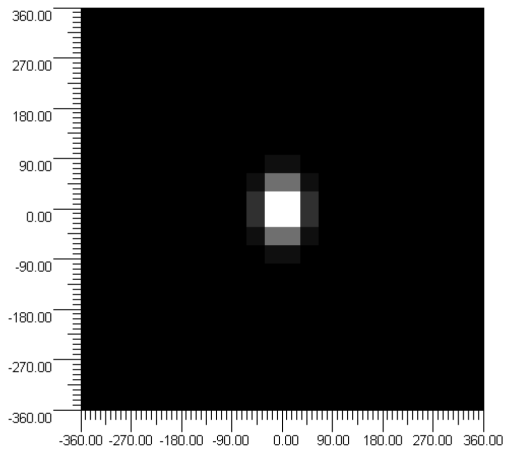

Fig.9 The near-field distribution restored by the 2nd iteration.

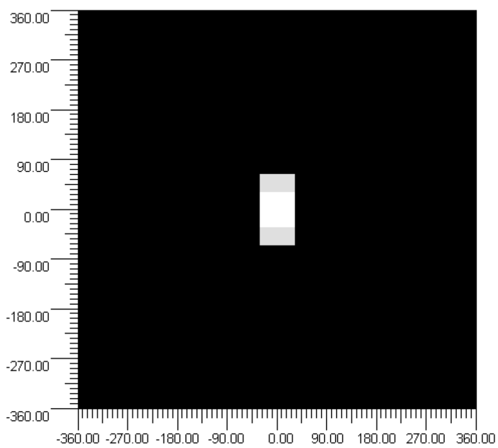

Fig.10 The near-field intensity distribution restored by the 3rd iteration.

Levels of iteration	Source apertures (in nanometers and wavelengths)	Profile error (in nanometers and wavelengths)	Notices
0 - iteration	Rectangle 90×180 (0.18×0.36) Rectangle 280×180 (0.56×0.36) (at 45°)	270 nm; 0.54 λ 140 nm; 0.28 λ	No extrapolation.
1st - iteration	Rectangle 90×180 Rectangle 280×180 (at 45°)	45 nm; 0.1 λ 70 nm; 0.14 λ	Short extrapolation by the 4th order polynomials.
2nd - iteration	Rectangle 90×180 Rectangle 280×180 (at 45°)	45 nm; 0.1 λ 56 nm; 0.11 λ	Short extrapolation by the 12th order polynomials.
3rd - iteration	Rectangle 90×180 Rectangle 280×180 (at 45°)	0.0 nm; 0.0 λ 28 nm; 0.06 λ	Long extrapolation by the 12th order polynomials.

7. REFERENCES

1. Chr. Obermüller and Kh. Karrai "Far-field characterization of diffracting circular apertures", *Appl. Phys. Lett.*, Vol. 67, No. 23, pp. 3408-3410, 1995.
2. Harris J.L. "Diffraction and Resolving Power", *J. Opt. Soc. Am.*, Vol. 54, No. 7, pp. 931-936, 1964.
3. Bates R.H.T. and McDonnell, *Image Restoration and Recognition*, Chs. 2-3, Oxford: University Press, 1986.
4. Stratton J.A., Morse P.M., Chu L.J., Little J.D.C., Corbato F.J., *Spheroidal wave functions*, N.Y.: John Wiley and Sons, Inc., 1956.
5. Voznessenski N.B. "Orthogonal polynomials to represent aberrations of optical systems with arbitrary kind of symmetry", *Izvestia VUZOV - USSR, Priborostroenie*, No. 5, 1982.
6. Rushforth C.K. and Harris R.W. "Restoration, Resolution and Noise", *J. Opt. Soc. Am.,* Vol. 58, No. 4, , pp. 539-545, 1968.
7. A.J. den Dekker and A. van den Bos, "Resolution: a survey", *J. Opt. Soc. Am.*, Vol. 14, No. 3, , pp. 547-557, 1997.