Three-dimensional total internal reflection microscopy

P. Scott Carney and John C. Schotland

Department of Electrical Engineering, Washington University, St. Louis, Missouri 63130

Received January 17, 2001

We investigate the inverse-scattering problem that arises in total internal reflection microscopy. An analytic solution to this problem within the weak-scattering approximation is used to develop a novel form of three-dimensional microscopy with subwavelength resolution. © 2001 Optical Society of America

OCIS codes: 290.3200, 180.6900, 100.3010.

There has been considerable recent interest in the development of near-field methods for optical microscopy.[1–3] These methods, which include near-field scanning optical microscopy[4] (NSOM) and total internal reflection microscopy (TIRM),[5,6] produce images with subwavelength resolution of effectively two-dimensional systems. It is now well recognized, however, that the interpretation of images of three-dimensional scattering objects is often problematic.[7–9] This difficulty has led to the use of inverse-scattering theory to elucidate the precise manner in which three-dimensional structure is encoded in the optical near field. In one recent study an analytic solution to the linearized near-field inverse-scattering problem for inhomogeneous media was employed to develop an image reconstruction algorithm for three-dimensional NSOM.[10] Additional progress toward the development of three-dimensional near-field microscopy has also been made in the case of TIRM, in which the three-dimensional information content of weakly scattered fields was investigated and employed to establish an inverse-scattering method to reconstruct the longitudinal structure of samples that are transversely homogeneous or at least separable.[11,12]

In this Letter we present the theoretical foundation of three-dimensional TIRM. By adapting the approach developed for three-dimensional NSOM, we show that it is possible to extract the full three-dimensional structure of a scattering object with subwavelength resolution. The proposed method obtains from an analysis of the inverse-scattering problem in which evanescent waves serve as a source of illumination, thereby exploiting the superoscillatory properties of such waves to encode structure on subwavelength scales. Our main result is an explicit inversion formula that leads to an image reconstruction algorithm that is computationally efficient and remarkably stable in the presence of noise. Computer simulations are used to illustrate our approach in model systems.

We begin by considering an experiment in which a monochromatic field is incident on a dielectric medium with complex susceptibility $\eta(\mathbf{r})$. For simplicity, we ignore the effects of polarization and consider the case of a scalar field $U(\mathbf{r})$ that obeys the reduced wave equation

$$\nabla^2 U(\mathbf{r}) + {k_0}^2 U(\mathbf{r}) = -4\pi {k_0}^2 \eta(\mathbf{r}) U(\mathbf{r}), \tag{1}$$

where k_0 is the free-space wave number. Following standard procedures, we find that $U(\mathbf{r})$ may be expressed as the integral equation

$$U(\mathbf{r}) = U^{(i)}(\mathbf{r}) + {k_0}^2 \int \mathrm{d}^3 r' G(\mathbf{r}, \mathbf{r}') U(\mathbf{r}') \eta(\mathbf{r}'), \tag{2}$$

where the outgoing Green's function $G(\mathbf{r}, \mathbf{r}')$ is given by

$$G(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik_0|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}, \tag{3}$$

and $U^{(i)}(\mathbf{r})$ is the incident field. We restrict our attention to the weak-scattering approximation, which is particularly suitable for the investigation of subwavelength structures. Accordingly, the scattered field $U^{(s)}(\mathbf{r}) \equiv U(\mathbf{r}) - U^{(i)}(\mathbf{r})$ may be calculated perturbatively to lowest order in η, with the result that

$$U^{(s)}(\mathbf{r}) = {k_0}^2 \int \mathrm{d}^3 r' G(\mathbf{r}, \mathbf{r}') U^{(i)}(\mathbf{r}') \eta(\mathbf{r}'). \tag{4}$$

We consider the situation in which the incident wave is a unit-amplitude evanescent wave with $U^{(i)}(\mathbf{r}) = \exp(i\mathbf{k}_1 \cdot \mathbf{r})$; see Fig. 1. Here the complex wave vector $\mathbf{k}_1$ is of the form $\mathbf{k}_1 = (\mathbf{q}_1, k_z(\mathbf{q}_1))$, with transverse part $\mathbf{q}_1$ and $k_z(\mathbf{q}_1) = ({k_0}^2 - {\mathbf{q}_1}^2)^{1/2}$. If the evanescent wave is generated by a prism with index of refraction n, then $k_0 \leq |\mathbf{q}_j| \leq nk_0$. Such evanescent waves are superoscillatory in the transverse plane and decay exponentially in the z direction. In the far zone, the scattered field behaves as an outgoing homogeneous wave with wave vector $\mathbf{k}_2 = (\mathbf{q}_2, k_z(\mathbf{q}_2))$ parallel to $\mathbf{r}$, and $|\mathbf{q}_2| \leq k_0$. The scattered field has the asymptotic form

$$U^{(s)}(\mathbf{r}) \sim \frac{\exp(ik_0 r)}{r} A(\mathbf{q}_1, \mathbf{q}_2), \tag{5}$$

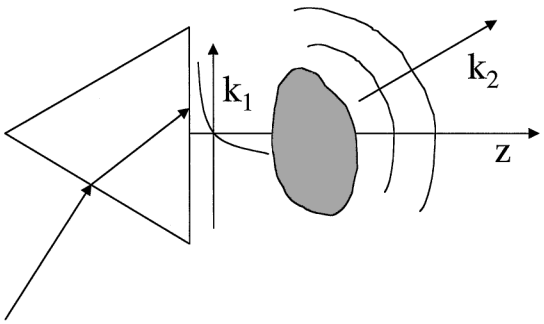

Fig. 1. Illustration of the measurement scheme. Evanescent waves are generated at the prism face by total internal reflection. The total internal reflection is then partly frustrated by the presence of the scatterer, scattering evanescent modes to the homogeneous mode that propagate to the far zone.

0146-9592/01/141072-03$15.00/0

where the scattering amplitude is given by the expression

$$A(\mathbf{q}_1, \mathbf{q}_2) = {k_0}^2 \int \mathrm{d}^3 r \exp\{i[\mathbf{k}_1(\mathbf{q}_1) - \mathbf{k}_2(\mathbf{q}_2)] \cdot \mathbf{r}\} \eta(\mathbf{r}), \tag{6}$$

where $\mathbf{k}_i(\mathbf{q}_i) \equiv \mathbf{k}_i$ for $i = 1, 2$ and the dependence of the wave vectors on their transverse parts has been made explicit.

In the inverse problem we wish to reconstruct $\eta(\mathbf{r})$ from measurements of $A(\mathbf{q}_1, \mathbf{q}_2)$ by solving integral equation (6) for $(\mathbf{q}_1, \mathbf{q}_2)$ in the data set $\mathcal{Q}$. Here $\mathcal{Q}$ specifies the available orientations of the incoming and outgoing waves. Let us rewrite Eq. (6) in the form

$$A(\mathbf{q}_1, \mathbf{q}_2) = {k_0}^2 \int \mathrm{d}^3 r \, K(\mathbf{q}_2, z; \mathbf{q}_1 - \mathbf{q}_2) \times \exp[i(\mathbf{q}_1 - \mathbf{q}_2) \cdot \boldsymbol{\rho}] \eta(\mathbf{r}), \tag{7}$$

where $\mathbf{r} = (\boldsymbol{\rho}, z)$ and

$$K(\mathbf{q}, z; \mathbf{Q}) = \exp\{i[k_z(\mathbf{Q} + \mathbf{q}) - k_z(\mathbf{q})]z\} \chi(\mathbf{q}, \mathbf{Q} + \mathbf{q}), \tag{8}$$

and $\chi(\mathbf{q}_1, \mathbf{q}_2)$ is unity if $(\mathbf{q}_1, \mathbf{q}_2) \in \mathcal{Q}$ and is zero otherwise. Equation (7) is similar in form to an equation that arises in the inverse scattering problem for three-dimensional NSOM.[10] Using the approach developed there, we may obtain an inversion formula for Eq. (6) by means of a functional singular-value decomposition, which leads to our main result:

$$\eta(\mathbf{r}) = \frac{1}{{k_0}^2 (2\pi)^2} \int \mathrm{d}^2\mathbf{Q} \exp(-i\mathbf{Q} \cdot \boldsymbol{\rho}) \int \mathrm{d}^2\mathbf{q} \mathrm{d}^2\mathbf{q}' \times K^*(\mathbf{q}, z; \mathbf{Q}) M^{-1}(\mathbf{q}, \mathbf{q}'; \mathbf{Q}) \chi(\mathbf{q}', \mathbf{Q} + \mathbf{q}') A(\mathbf{q}', \mathbf{Q} + \mathbf{q}'). \tag{9}$$

Here M^{-1} is obtained from M, M being given by the expression

$$M(\mathbf{q}, \mathbf{q}'; \mathbf{Q}) = \int_0^L K(\mathbf{q}, z; \mathbf{Q}) K^*(\mathbf{q}', z; \mathbf{Q}) \mathrm{d}z. \tag{10}$$

The parameter L in Eq. (10) denotes the range of $\eta(\mathbf{r})$ in the z direction and acts as a means of including some prior knowledge in the algorithm.

Regularization is achieved in a straightforward manner by conditioning of M, i.e., by elimination of the eigenfunctions associated with eigenvalues $\sigma(\mathbf{Q})$, such that $\sigma(\mathbf{Q}) < \epsilon \sigma_{\max}(\mathbf{Q})$, where $\sigma_{\max}(\mathbf{Q})$ is the largest eigenvalue for fixed $\mathbf{Q}$ and ϵ is chosen appropriately, given the level of noise. In principle, other approaches such as Tikhonov regularization can be employed.

To demonstrate the feasibility of the inversion we have numerically simulated the reconstruction of $\eta(\mathbf{r})$ for a collection of spherical scatterers. We calculated the forward data by considering the scattering of evanescent waves from a homogeneous sphere including multiple scattering by means of a partial wave expansion. For a sphere of radius a centered at $\mathbf{r}_0$ with refractive index n, n being related to the susceptibility by $n^2 = 1 + 4\pi\eta$, it may be found that

$$A(\mathbf{k}_1, \mathbf{k}_2) = \exp[i(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}_0] \sum_{l=0}^{\infty} (2l + 1) A_l P_l(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2), \tag{11}$$

where A_l are the usual partial wave expansion coefficients,[13] P_l are the Legendre polynomials, and the caret has the meaning $\hat{\mathbf{k}} = \mathbf{k}/\sqrt{\mathbf{k} \cdot \mathbf{k}^*}$. Since we consider evanescent waves, the argument of the Legendre polynomials in Eq. (11) may exceed unity. The series may nonetheless be shown to be convergent because of the rapid decay of A_l with increasing l.

The forward data were obtained for a collection of four spheres of radius $\lambda/20$. All scatterers are present simultaneously in the forward simulation, and intersphere scattering is neglected. The incident evanescent waves were taken to have transverse wave vectors $\mathbf{q}_1$ on a Cartesian grid with spacing $k_0/4$ such that $k_0 \leq |\mathbf{q}_1| \leq n k_0$, n being the refractive index of the prism used to generate the incident waves. Values of $\mathbf{q}_2$ were also taken on the grid with spacing $k_0/4$ and $|\mathbf{q}_2| \leq k_0$, consistent with a measurement scheme in which the scattered field is measured in the far zone of the lower half-space. The spheres were arranged in two layers, one equatorial plane coincident with the $z = \lambda/20$ plane, the other with the $z = \lambda/4$ plane. In each layer, one sphere was taken to have susceptibility $4\pi\eta = 0.44$ (index $n = 1.2$) and one sphere was taken to have susceptibility $4\pi\eta = 0.4 + 0.48i$ (index $n = 1.2 + 0.2i$). In each of the simulations complex Gaussian noise of zero mean was added to the signal at various levels as indicated. Simulations were performed for two different prisms, one (Fig. 2) with an index of $n = 10$, as might be encountered in the infrared, and another (Fig. 3) with an index of $n = 4$, as might be encountered in the visible. The data sets thus consisted of 4980×49 points for the higher-index prism and 752×49 data points for the lower-index prism. The regularization parameter ϵ was taken to be 10^{-6} in computing the tomograph at $z = \lambda/20$ and 10^{-5} at $z = \lambda/4$.

One can see from the reconstructions that the real and imaginary parts of the susceptibility may be found separately and that the reconstructions are subwavelength resolved. The resolution depends both on the size of the regularization parameter, which indirectly sets the number of singular functions used in the reconstruction, and on the depth, a consequence of the fact that the probe fields decay exponentially into the sample, resulting in the loss of high-frequency Fourier components of the spatial structure of the susceptibility. The tomographs at the $z = \lambda/20$ layer are more highly resolved for the higher-index prism than for the lower-index prism, but there is little difference at the $z = \lambda/4$ layer.

In conclusion, we have developed an analytic solution to the linearized inverse-scattering problem that arises in three-dimensional TIRM. Using this result, we have shown that it is possible to reconstruct the three-dimensional structure of a scattering medium

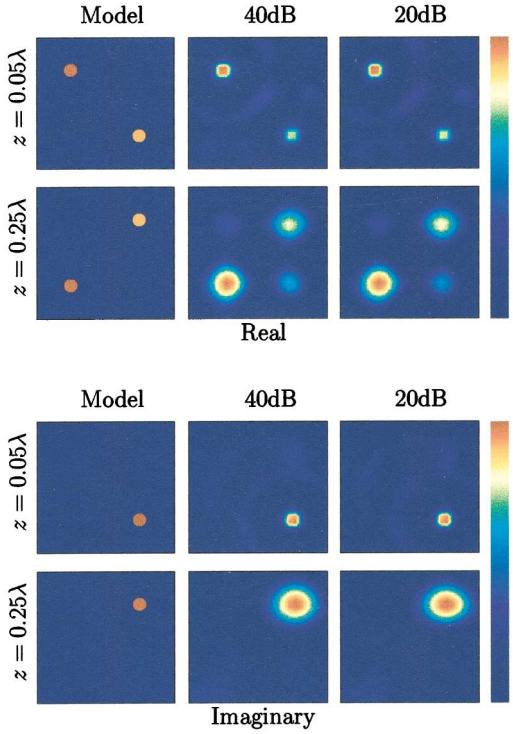

Fig. 2. Reconstructed tomographs of the real and imaginary parts of the susceptibility with a prism of refractive index $n = 10$. The signal-to-noise ratio is given in decibels above each column. The images were plotted with the linear color scale indicated to the right. The field of view in each image is $\lambda \times \lambda$.

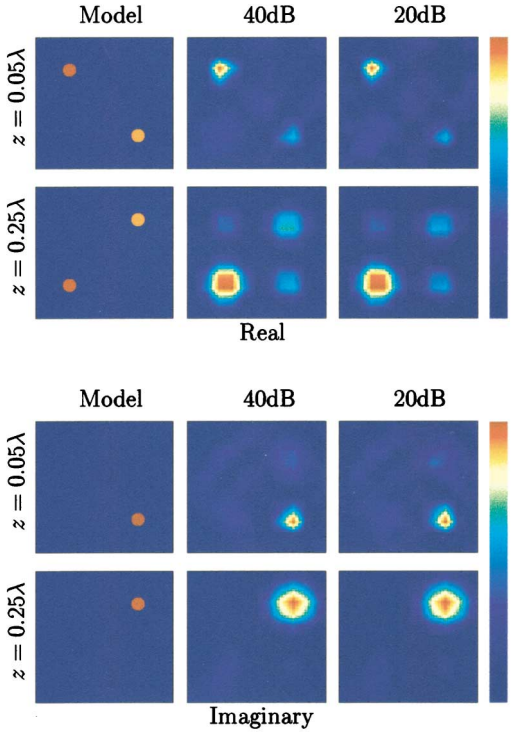

Fig. 3. Reconstructed tomographs of the real and imaginary parts of the susceptibility with a prism of refractive index $n = 4$. All other parameters are as indicated in Fig. 2.

on subwavelength scales. The improved resolution is made possible by the use of evanescent waves as illumination, so that the far field of the resulting scattered wave contains information about the scattering medium on subwavelength scales. Subwavelength resolution may be achieved from far-zone measurements because the illuminating field is in the near zone of the sample, a consequence of generalized reciprocity.[14] That is, even though the measured fields are homogeneous, and therefore spatially band limited, the illumination field is not.

J. C. Schotland's e-mail address is jcs@ee.wustl.edu.

References

1. M. A. Paesler and P. J. Moyer, *Near Field Optics* (Wiley, New York, 1996).
2. D. Courjon and C. Bainier, Rep. Prog. Phys. **57,** 989 (1994).
3. C. Girard and A. Dereux, Rep. Prog. Phys. **59,** 657 (1996).
4. E. Betzig and J. K. Trautman, Science **257,** 189 (1992).
5. C. W. McCutchen, Rev. Sci. Instrum. **35,** 1340 (1964).
6. P. A. Temple, Appl. Opt. **20,** 2656 (1981).
7. D. Van Labeke and D. Barchiesi, J. Opt. Soc. Am. A **9,** 732 (1992).
8. R. Carminati and J. Greffet, Opt. Commun. **116,** 316 (1995).
9. R. Carminati and J. Greffet, J. Opt. Soc. Am. A **12,** 2716 (1995).
10. P. S. Carney and J. C. Schotland, Appl. Phys. Lett. **77,** 2798 (2000).
11. D. G. Fischer, J. Mod. Opt. **47,** 1359 (2000).
12. D. G. Fischer, Opt. Lett. **25,** 1529 (2000).
13. The coefficient A_l is given by the expression $A_l = i\beta_l/[k_0(\beta_l - i\gamma_l)]$ with $\beta_l = j_l(nk_0a)j_l'(k_0a) - nj_l'(nk_0a)j_l(k_0a)$ and $\gamma_l = nj_l'(nk_0a)n_l(k_0a) - j_l(nk_0 \times a)n_l'(k_0a)$.
14. R. Carminati, M. Nieto-Vesperinas, and J. Greffet, J. Opt. Soc. Am. A **15,** 706 (1998).