Lateral resolution enhancement with standing evanescent waves

George E. Cragg and Peter T. C. So

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received August 10, 1999

A high-resolution fluorescence microscopy technique has been developed that achieves a lateral resolution of better than one sixth of the emission wavelength (FWHM). By use of a total-internal-reflection geometry, standing evanescent waves are generated that spatially modulate the excitation of the sample. An enhanced two-dimensional image is formed from a weighted sum of images taken at different phases and directions of the standing wave. The performance of such a system is examined through theoretical calculations of both the point-spread function and the optical transfer function. © 2000 Optical Society of America

OCIS codes: 170.0180, 180.2520, 180.3170, 180.5810.

The Rayleigh criterion quantifies the resolving capabilities of an incoherent imaging system as the FWHM of the point-spread function (PSF). By this criterion, the lateral resolution (Δr) of conventional microscopy is given approximately by $\Delta r \approx \lambda_e/2\mathrm{NA} \approx \lambda_e/2$, where λ_e is the emission wavelength and NA is the numerical aperture of the imaging optics. Many techniques such as confocal and 4 pi microscopy have attained $\Delta r \approx \lambda_e/3$ (Ref. 1; p. 48, Ref. 2). Additionally, atomic scale resolution has been achieved in semiconductor samples with scanning probe techniques such as near-field scanning optical microscopy, scanning tunnel microscopy, and atomic-force microscopy.[3,4] However, not only do these scanning probe techniques suffer from slow imaging speed in general but they also have limited applicability in soft biological specimens. We have developed standing-wave total-internal-reflection fluorescence (SWTIRF) imaging, an all-optical technique that can achieve a lateral resolution of better than $\lambda_e/6$.

Axial subsectioning in thin specimens has been demonstrated by use of standing-wave excitation to produce a high-frequency intensity modulation along the optical axis.[5,6] With a total-internal-reflection[7] geometry, a similar intensity pattern can be generated laterally across a sample. In standard total-internal-reflection microscopy, laser light is totally reflected off the surface of a prism to form an exponentially decaying evanescent wave that is confined to approximately 50–100 nm in the lower index medium. The evanescent field then excites a fluorescently labeled specimen, thus allowing an image to be formed. Retroreflecting an s-polarized excitation beam back onto itself generates standing evanescent waves of wavelength $\lambda/n\sin(\theta_i)$ on the surface, where λ is the wavelength of the excitation, n is the refractive index of the prism, and θ_i is the angle of incidence of the incoming light.[7,8] Because the excitation is spatially modulated by the high-frequency standing wave, the intensity distribution on the sample plane, $I_{\mathrm{sp}}(\mathbf{r})$, is given by

$$I_{\mathrm{sp}}(\mathbf{r}) = O(\mathbf{r})[1 + \cos(Kx + \phi)], \tag{1}$$

$$K = 4\pi n \sin(\theta_i)/\lambda, \tag{2}$$

where the modulation is taken along the x axis and ϕ is an unknown phase that specifies the position of the standing wave relative to the object function, $O(\mathbf{r})$. Assuming a magnification of unity throughout, the image, $I_0(\mathbf{r})$, is given by the convolution of the excitation pattern with the PSF of conventional microscopy (p. 18, Ref. 2), $P(\mathbf{r})$:

$$I_0(\mathbf{r}) = O(\mathbf{r})[1 + \cos(Kx + \phi)] \otimes P(\mathbf{r}), \tag{3}$$

where $\otimes$ denotes the convolution operation.

To illustrate a direct way to achieve resolution improvement from a sum of appropriately weighted standing-wave images, the method outlined by Krishnamurthi *et al.*[6] is recast in real space. In the analysis, three images are required, each taken at a different standing-wave phase from the other two. All three can be compactly written in a single expression, $I_j(\mathbf{r})$:

$$I_j(\mathbf{r}) = O(\mathbf{r}) \otimes P(\mathbf{r}) + \{(1/2)\exp[i(Kx + \phi + j\pi/2)]O(\mathbf{r})\} \otimes P(\mathbf{r}) + \{(1/2)\exp[-i(Kx + \phi + j\pi/2)]O(\mathbf{r})\} \otimes P(\mathbf{r}), \qquad j = 0, 1, 2, \tag{4}$$

where the phases of ϕ, $\phi + \pi/2$, and $\phi + \pi$ are chosen. These images can be algebraically manipulated:

$$O(\mathbf{r}) \otimes P(\mathbf{r}) = (1/2)[I_0(\mathbf{r}) + I_2(\mathbf{r})], \tag{5}$$

$$\{(1/2)\exp[i(Kx + \phi)]O(\mathbf{r})\} \otimes P(\mathbf{r}) = (1/4)[(1 - i)I_0(\mathbf{r}) + 2iI_1(\mathbf{r}) - (1 + i)I_2(\mathbf{r})], \tag{6}$$

$$\{(1/2)\exp[-i(Kx + \phi)]O(\mathbf{r})\} \otimes P(\mathbf{r}) = (1/4)[(1 + i)I_0(\mathbf{r}) - 2iI_1(\mathbf{r}) + (-1 + i)I_2(\mathbf{r})]. \tag{7}$$

Using the identity

$$\{\exp[\pm i(Kx + \phi)]O(\mathbf{r})\} \otimes P(\mathbf{r}) = \exp[\pm i(Kx + \phi)]\{O(\mathbf{r}) \otimes [\exp(\mp iKx)P(\mathbf{r})]\}, \tag{8}$$

0146-9592/00/010046-03$15.00/0

we can transfer the modulation from the object function to the conventional PSF by first multiplying Eqs. (6) and (7) by $\exp[-i(Kx + \alpha)]$ and $\exp[i(Kx + \alpha)]$, respectively, and then summing the results with Eq. (5), thus yielding an enhanced image, $I(\mathbf{r})$:

$$\begin{aligned} I(\mathbf{r}) &= O(\mathbf{r}) \otimes \{P(\mathbf{r})[1 + \cos(Kx - \phi + \alpha)]\} \\ &= I_0(\mathbf{r})\,(1/2)[1 + \sqrt{2}\cos(Kx + \alpha + \pi/4)] \\ &\quad + I_1(\mathbf{r})\sin(Kx + \alpha) + I_2(\mathbf{r})\,(1/2) \\ &\quad \times [1 - \sqrt{2}\cos(Kx + \alpha - \pi/4)]. \end{aligned} \tag{9}$$

Hence, $I(\mathbf{r})$ is the sum of the three images of Eq. (4), each receiving its own sinusoidal weighting factor with arbitrary phase α.

As the optimum resolution is achieved when $\alpha = \phi$, α must be properly chosen in the weighting factors. This choice can be made in either of two main ways. First, on integrating the left-hand side of Eq. (9) over the detection area, one can see that the total intensity is maximized for $\alpha = \phi$. Therefore, α can be adjusted until this maximum intensity image is obtained. Alternatively, a point source on the prism would serve as a reference against which ϕ could be measured.

With ϕ determined, the PSF of the SWTIRF scan along the x axis, $P_{\mathrm{SWx}}(\mathbf{r})$, is given as

$$P_{\mathrm{SWx}}(\mathbf{r}) = [1 + \cos(Kx)]P(\mathbf{r}). \tag{10}$$

Since the diffraction-limited PSF is modulated by the $1 + \cos(Kx)$ term of the intensity distribution, the PSF is narrowed along the x axis while its conventional width along the y axis is retained [Fig. 1(b)]. Note that the degree of enhancement along x depends on the spatial frequency K, which, in turn, depends linearly on the refractive index of the prism, n. Because the periodicity increases with n, high-index materials will provide the greatest narrowing of the PSF's central maximum. As is demonstrated in Fig. 1(a), the FWHM (1D SWTIRF) along x is approximately $\lambda_e/6$ and $\lambda_e/9$ for quartz ($n = 1.46$ at 532 nm) and lithium niobate ($LiNbO_3$, n = ordinary index = 2.32 at 532 nm) prisms, respectively. However, as n is increased, the sidebands in the PSF increase as well. Fortunately, it may be possible to remove the sidebands by use of simple deconvolution algorithms.[9] Additionally, SWTIRF may be combined with two-photon[10] or stimulated-emission[11–13] imaging techniques for further resolution enhancement as well as sideband suppression.

An alternative way to describe the performance of an optical system is to examine its optical transfer function (OTF), which, as the Fourier transform of the PSF, reveals both the spatial frequencies that are present in the image and the relative weight of these frequency components. Because the OTF can be used to describe the low-pass filtering characteristics of the imaging system, the resolution is determined from the cutoff frequency of the OTF. The OTF of the one-dimensional SWTIRF application, $\tilde{P}_{\mathrm{SWx}}(\mathbf{k})$ [Fig. 2(a)], is given as the normalized Fourier transform of Eq. (10):

$$\begin{aligned} \tilde{P}_{\mathrm{SWx}}(\mathbf{k}) &= [2 + \tilde{P}(-K\mathbf{e}_x) + \tilde{P}(K\mathbf{e}_x)]^{-1} \\ &\quad \times [2\tilde{P}(\mathbf{k}) + \tilde{P}(\mathbf{k} - K\mathbf{e}_x) + \tilde{P}(\mathbf{k} + K\mathbf{e}_x)], \end{aligned} \tag{11}$$

where $\tilde{P}(\mathbf{k})$ is the incoherent OTF of conventional microscopy (p. 35, Ref. 2; Ref. 14) and K is defined by Eq. (2). Therefore, the $\tilde{P}_{\mathrm{SWx}}(\mathbf{k})$ function consists of a sum of three $\tilde{P}(\mathbf{k})$ functions, one at the center, one shifted by K along k_x, and the third shifted by $-K$ along k_x, where the center function has twice the weight of the outer two. If K is large enough to create frequency gaps along k_x ($K \geq 8\pi\mathrm{NA}/\lambda_e$), the outer functions will separate from the center one, thereby causing severe ringing in the PSF. For cases such as those for quartz and $LiNbO_3$ prisms, in which there are no frequency gaps, the cutoff frequency of the one-dimensional SWTIRF is $4\pi\mathrm{NA}/\lambda_e + K$ along the k_x axis but retains the conventional value of $4\pi\mathrm{NA}/\lambda_e$ along the k_y axis.

Obtaining resolution enhancement in two dimensions requires multiple images taken along different standing-wave directions in the sample plane. For the

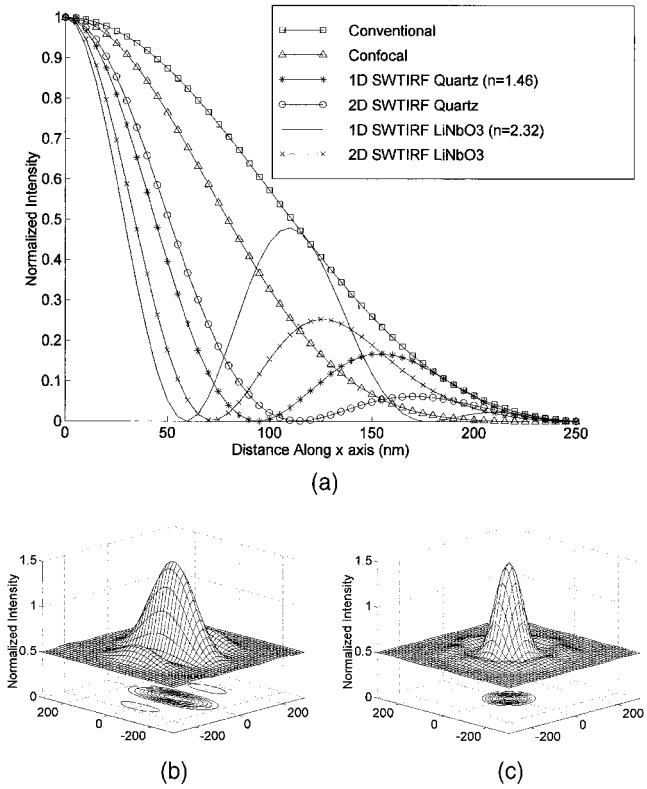

Fig. 1. For the representative parameters of $N = 4$, $\theta_i = 75°$, NA = 1.3, $\lambda = 532$ nm, and $\lambda_e = 560$ nm, plots are generated that show (a) the PSF's of conventional, confocal, and SWTIRF microscopies, (b) the PSF of a one-dimensional (1D) SWTIRF (quartz, $n = 1.46$) scan [Eq. (10)], (c) the symmetric PSF of two-dimensional (2D) SWTIRF (quartz) generated by Eq. (14).

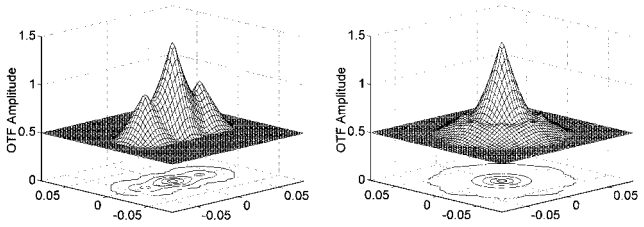

Fig. 2. (a) and (b), Frequency space counterparts (OTF's) of Figs. 1(b) and 1(c), respectively.

case in which the modulation is at some angle ξ with respect to the x axis, Eq. (10) may be rewritten as

$$P_{\mathrm{SW}\xi}(\mathbf{r},\xi) = \{1 + \cos[Kr\cos(\theta - \xi)]\}P(\mathbf{r}), \qquad (12)$$

where $\mathbf{r} = (\mathbf{r},\theta)$. In constructing the enhanced two-dimensional image we impose several constraints on the PSF and the OTF. First, the cutoff frequency of the OTF must approach $4\pi\mathrm{NA}/\lambda_e + K$ in all directions, thereby requiring a total of $N > 2$ scans to be performed, one at every π/N rad between 0 and π in the sample plane. Consistency with incoherent imaging requires the enhanced image to be linear in intensity (p. 33, Ref. 2). Hence these N scans are added together, with all scans weighted equally and no direction favored over any other. As this summing operation may cause the lower-frequency components ($k < 4\pi\mathrm{NA}/\lambda_e$) to dominate the OTF, it may be necessary to subtract an appropriately weighted diffraction-limited image. Moreover, it is required that the PSF be nonnegative and decay to zero for some $r < \lambda_e/2$. Although the nonnegativity requirement may not be essential, we introduce it to conform to conventional optics. A linear combination that realizes these considerations is, with the help of Eq. (12), given in the two-dimensional SWTIRF PSF, $P_{\mathrm{SW}}(\mathbf{r}, N)$:

$$P_{\mathrm{SW}}(\mathbf{r}, N) = \frac{1}{1-\beta}\left(\frac{1}{N}\sum_{m=0}^{N-1}\{1 + \cos[Kr\cos(\theta - m\pi/N)]\} - \beta\right)P(\mathbf{r}) = \frac{1}{1-\beta}S(\mathbf{r},N)P(\mathbf{r}), \quad (13)$$

where $S(\mathbf{r}, N)$ is the term in parentheses, $1/(1-\beta)$ is the normalization factor, and β is a constant that determines the weight of the diffraction-limited image. In this case, the enhanced resolution is achieved through the symmetric modulation $S(\mathbf{r},N)$. By the Rayleigh criterion, β is chosen to minimize the radial distance between the center and the first minimum of $S(\mathbf{r},N)$. In restricting $S(\mathbf{r},N)$ to nonnegative values, we set this minimum to zero. It is therefore required that $S(\mathbf{r},N)$ and its first derivative with respect to r vanish for some distance $r_{\min} < \lambda_e/2$, thereby obtaining the constant $\beta = 0.3$. On substituting this value into Eq. (13) we have

$$P_{\mathrm{SW}}(\mathbf{r},N) = 1.43\left(\frac{1}{N}\sum_{m=0}^{N-1}\{1 + \cos[Kr\cos(\theta - m\pi/N)]\} - 0.3\right)P(\mathbf{r}); \qquad (14)$$

Fig. 1(c) shows the case for $N = 4$. The reciprocal space counterpart of Eq. (14) is given as $\tilde{P}_{\mathrm{SW}}(\mathbf{k}, N)$ [Fig. 2(b)]:

$$\tilde{P}_{\mathrm{SW}}(\mathbf{k},N) = \gamma\left(\frac{1}{N}\sum_{m=0}^{2N-1}\tilde{P}\{\mathbf{k} + K[\cos(m\pi/N)\mathbf{e}_x + \sin(m\pi/N)\mathbf{e}_y]\} + 1.7\tilde{P}(\mathbf{k})\right), \qquad (15)$$

where γ is a normalization factor. Whereas adding scans together in this way produces a symmetric PSF with reduced sidebands, the resultant PSF is slightly broader than the x-axis profile of $P_{\mathrm{SWx}}(\mathbf{r})$. From Fig. 1(a) it can be seen that $P_{\mathrm{SW}}(\mathbf{r})$ has a FWHM of approximately $\lambda_e/5$ and $\lambda_e/8$ for quartz and $LiNbO_3$ prisms, respectively.

Finally, the range of applicability of SWTIRF must be recognized. Because of the shallow depth of field, this technique will find applications primarily in the study of surface phenomena, such as in cell membranes. Moreover, small vertical displacements of the specimen will cause significant variations in intensity. However, no mechanical contact with the specimen is required, because the evanescent wave itself acts as the probing mechanism. This would permit high-fidelity imaging of soft biological samples. This technique may make possible video rate image acquisition, as not more than three images are required for each standing-wave direction.

This research was supported by the National Science Foundation (research grant MCB-9604382). G. E. Cragg's e-mail address is cragg@mit.edu.

References

1. M. Schrader, S. W. Hell, and H. T. M. van der Voort, J. Appl. Phys. **84,** 4033 (1998).
2. T. Wilson and C. Sheppard, *Theory and Practice of Scanning Optical Microscopy* (Academic, New York, 1984).
3. E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251,** 1468 (1991).
4. G. Binnig, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56,** 930 (1986).
5. B. Bailey, D. L. Farkas, D. L. Taylor, and F. Lanni, Nature **366,** 44 (1993).
6. V. Krishnamurthi, B. Bailey, and F. Lanni, Proc. SPIE **2655,** 18 (1996).
7. D. Axelrod, Annu. Rev. Biophys. Bioeng. **13,** 247 (1984).
8. J. R. Abney, B. A. Scalettar, and N. L. Thompson, Biophys. J. **61,** 542 (1992).
9. P. E. Hanninen, S. W. Hell, J. Salo, and E. Soini, Appl. Phys. Lett. **66,** 1698 (1995).
10. W. Denk, J. H. Strickler, and W. W. Webb, Science **248,** 73 (1990).
11. S. W. Hell and J. Wichmann, Opt. Lett. **19,** 780 (1994).
12. C. Y. Dong, P. T. C. So, T. French, and E. Gratton, Biophys. J. **69,** 2234 (1995).
13. T. A. Klar and S. W. Hell, Opt. Lett. **24,** 954 (1999).
14. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996), p. 143.