NANO LETTERS

2006
Vol. 6, No. 11
2609−2616

Differential Near-Field Scanning Optical Microscopy

Aydogan Ozcan,*,† Ertugrul Cubukcu,‡,§ Alberto Bilenca,†,§ Kenneth B. Crozier,‡ Brett E. Bouma,† Federico Capasso,‡ and Guillermo J. Tearney†

Wellman Center for Photomedicine, Harvard Medical School, Boston, Massachusetts 02114, and Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

Received September 7, 2006; Revised Manuscript Received October 6, 2006

ABSTRACT

We theoretically and experimentally illustrate a new apertured near-field scanning optical microscopy (NSOM) technique, termed differential NSOM (DNSOM). It involves scanning a relatively large (e.g., 0.3–2 μm wide) rectangular aperture (or a detector) in the near-field of an object and recording detected power as a function of the scanning position. The image reconstruction is achieved by taking a two-dimensional derivative of the recorded power map. Unlike conventional apertured NSOM, the size of the rectangular aperture/detector does not determine the resolution in DNSOM; instead, the resolution is practically determined by the sharpness of the corners of the rectangular aperture/detector. Principles of DNSOM can also be extended to other aperture/detector geometries such as triangles and parallelograms.

Near-field scanning optical microscopy[1−13] (NSOM) is an exciting imaging modality which permits super-resolution imaging of samples, breaking the diffraction barrier of light. In conventional aperture-type near-field scanning optical microscopy, the resolution is limited by the aperture size of the tip.[1−9] However since the effective transmission area decreases as the fourth power of the aperture diameter,[14−16] improved resolution comes at the price of a sharp decrease in signal-to-noise ratio (SNR) and contrast of the NSOM image. Several techniques to improve the light throughput of NSOM tips have been recently proposed;[1] however, these approaches either are difficult to fabricate or cause a reduction in contrast to improve the light transmission. Furthermore, the effective diameter of the aperture in conventional aperture-type NSOM cannot be physically smaller than twice the skin depth of the metal tip, which implies a lower bound of ≈20 nm for the effective aperture width.[1] To solve some of these problems, scattering-type apertureless NSOM has also been successfully demonstrated as a promising near-field imaging modality.[1,17−21]

In this work, we introduce an alternative new approach for aperture-type NSOM, termed differential near-field scanning optical microscopy (DNSOM), which involves scanning a rectangular (e.g., a square) aperture (or a detector) in the near-field of the object of interest and recording the power of the light collected from the rectangular structure as a function of the scanning position. The image reconstruction is achieved by taking a two-dimensional (2D) derivative of the recorded power map. In contrast with conventional NSOM, here the size of the rectangular aperture or the detector does not determine the resolution of the recovered image. In DNSOM, the resolution is instead determined by the sharpness of the corners of the rectangle and the step size of the scan.

To explain the principles of DNSOM, we model the optical power transmissivity of the 2D object of interest by $O(x,y)$, where x and y denote the coordinates in the plane of the sample.[22] The same derivation could also be extended to a reflection geometry rather than transmission. For this derivation we will assume the object to be infinitely thin. This assumption is also made in other apertured or apertureless NSOM approaches making NSOM primarily a 2D imaging modality. However, recently there has been some interesting work on extending near-field microscopy to all three dimensions by treating the object recovery as an inverse scattering problem.[23,24] For our derivation, let us further assume that, without loss of generality, the DNSOM aperture is a square with a width of W, i.e., the power transmissivity of the square aperture is given by

$$\mathrm{Rect}(x,y) = \begin{cases} 1, & \text{if } W/2 > x > -W/2 \text{ and } W/2 > y > -W/2 \\ 0, & \text{elsewhere} \end{cases}$$

In this formulation, the effect of the skin depth at the walls of the square aperture has been ignored, an omission that will be addressed in the discussion to follow. Here, we should emphasize that this same analysis could also be extended to a rectangular DNSOM aperture, where in general the width

* Corresponding author. E-mail: aozcan@mgh.harvard.edu.
† Wellman Center for Photomedicine, Harvard Medical School.
‡ Division of Engineering and Applied Sciences, Harvard University.
§ Contribution of these two authors is equal.

10.1021/nl062110v CCC: $33.50
Published on Web 10/20/2006

and height of the aperture are different. For a scanning step size of Δx and Δy along x and y, respectively, the detected power in the far-field, collected from the square aperture, as a function of the scanning coordinates, $m\Delta x, n\Delta y$ (where m and n are integers) can be written as

$$P(m\cdot\Delta x, n\cdot\Delta y) = \int\int O(x,y)\cdot \mathrm{Rect}(x - m\Delta x, y - n\Delta y)\, dx\, dy \quad (1)$$

In eq 1, nonpropagating evanescent waves that make up the high spatial frequencies of $O(x,y)$ are assumed to reach the far-field detector via scattering from the edges of the square. In other words, extremely high spatial frequencies corresponding to the corners of the square aperture shift the nonpropagating high spatial frequency band of $O(x,y)$ into the passband of the numerical aperture of the far-field detection system. A similar assumption is also made for the operational principle of conventional NSOM, where the power detection is also made in the far-field.[1] In addition to scanning a large aperture over the sample, this derivation also holds for scanning a square-area *detector* in the sample's near field, which is advantageous because it eliminates losses and spatial frequency distortions associated with information transfer from the near field into the far field.[1] Furthermore, similar to conventional NSOM, we also assume in eq 1 that the near-field pattern of the object function, $O(x,y)$, is altered by the presence of the aperture/detector area in the same way for all the scanning positions, the overall effect of which is minimized by the 2D derivative operation.

By a simple change of variables, $x' = m\Delta x$ and $y' = n\Delta y$, one can take the 2D derivative of eq 1, and after some algebraic steps arrive at

$$\frac{\partial^2 P(x',y')}{\partial x'\cdot\partial y'} = O(x' - W/2, y' - W/2) + O(x' + W/2, y' + W/2) - O(x' + W/2, y' - W/2) - O(x' - W/2, y' + W/2) \quad (2)$$

Equation 2 indicates that by taking a 2D derivative of the scanning power output of the square aperture/detector of DNSOM, four replicas of the original object function $O(x,y)$ are recovered. Each replica is centered around one of the corners of the square.

Therefore, theoretically, the resolution of DNSOM recovery depends on (1) the scanning step size and (2) the sharpness of the corners of the square. Since the scanning step size can be made to be less than a nanometer with the current state-of-the-art piezoelectric scanners, we can state that the resolution of DNSOM practically depends only on the sharpness of the corners of the square structure. In the derivation of eq 2, corners of the square are assumed to be ideal as defined by $\mathrm{Rect}(x,y)$; i.e., after 2D differentiation each corner yields a 2D Dirac-delta function, $\delta(x,y)$. For an imperfect square with slightly round corners (where $\mathrm{Rect}(x,y)$ can no longer perfectly define the DNSOM aperture), after 2D differentiation each corner of the square will yield a point-spread function, $p(x,y)$, where $p(x,y) \neq \delta(x,y)$. Each replica image will therefore be equal to the convolution of the true object function with the point spread function, i.e., $O(x,y) \otimes p(x,y)$, where "$\otimes$" denotes the spatial convolution operation. Since the $\partial^2/\partial x'\,\partial y'$ operator creates a narrower point-spread function than the actual geometrical corner roundness, the use of focused-ion-beam milling or electron-beam writing could, in principle, result in a DNSOM point-spread function narrower than $\sim$10 nm.[1] A similar discussion also applies to the skin depth (L_s) of the optical field at the square edges. In conventional NSOM, the lower bound of the effective aperture diameter is roughly $2L_s$, whereas in DNSOM, the limiting effect of the skin depth is reduced to $\sim L_s$, since each of the side walls of the square operates separately in DNSOM.

The width (W) of the square aperture/detector in DNSOM affects two quantities: (1) the maximum area of the object that can be imaged and (2) the light throughput. The field of view (FOV) area for DNSOM with a square aperture is $2W \times 2W = 4W^2$. In order to avoid irreparable information loss, the object should be smaller than $2W$ in either dimension. For flat samples, a larger FOV can be achieved using, in parallel to the scanning DNSOM aperture/detector, a movable rectangular mask that has an area of $2W \times 2W$. This way, by translating the mask by $2W$ along x and/or y, new regions on the large object surface can be scanned using DNSOM, without spatial information loss.

In terms of light throughput, there is an exponential penalty as the width is reduced in the region $W \ll \lambda$, where λ is wavelength of illumination light. For the other limiting case of $W \gg \lambda$, the transmission of the square aperture increases as $\sim W^2$. In between these two regions ($W \sim \lambda$), the ratio of the effective transmission area to the real physical area of a small hole approaches unity and the light throughput penalty is minimized.[14−16] Selection of the optimal aperture size is therefore a function of object size and SNR considerations. If the maximum object dimensions are smaller than the wavelength of light, an aperture size $W \sim \lambda$ should be chosen for high light throughput. Making the aperture larger in this case would only increase the shot-noise and not provide additional information on the sample. If the object size is larger than the wavelength of light, then W should be approximately half the largest dimension of the object. The upper bound of W is governed by the *dynamic range* of the detection system and the *sensitivity* required for accurate DNSOM image recovery.

We first illustrate DNSOM with a numerical example. For this purpose, we created an arbitrary digital object (a microscope image of a tissue sample containing an overlaid character "v") as shown in Figure 1a, where the value of each point on the object varied between 0 and 1. We modeled this object such that each pixel on the image corresponded to 10 nm (i.e., $\Delta x = \Delta y = 10$ nm), and the total extent of the image was 2 μm $\times$ 2 μm. In the simulation, a square DNSOM aperture of 2 μm $\times$ 2 μm was scanned over the sample surface, while assuming a perfect plane wave illumination. Using eq 1, the scanning output power was computed as shown in Figure 1b. In this numerical simulation, the DNSOM aperture was assumed to be positioned

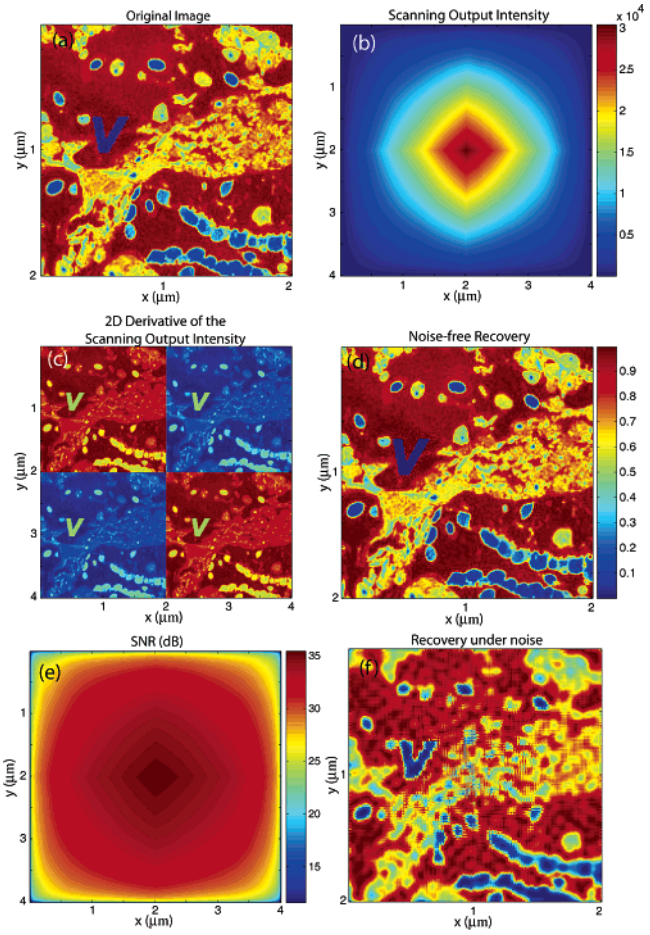

Figure 1. Results from a numerical simulation to illustrate the principles of DNSOM. A square DNSOM aperture of 2 μm × 2 μm was assumed and the simulation scanned this aperture exactly over the sample surface. (a) The original image (a microscope image of a tissue sample with an overlaid "v") that was arbitrarily chosen for this numerical example. Each pixel on the image was simulated to be 10 nm, and the total extent of the image was 2 μm × 2 μm. (b) The resultant DNSOM scanning output intensity. (c) 2D derivative of Figure 1b. Red and blue colors represent positive and negative values, respectively. As expected from our theory, two of the corners (first and fourth) provided the original replica of the image where as the other two corners (second and third) resulted in a negative inverted version of the original image. (d) The final recovery, which was obtained by tiling the recovered images of Figure 1c. (e) To simulate the noise performance of DNSOM, random Gaussian noise was added at each pixel of the scanning output and the resultant SNR at each point of the scanning is shown. (f) DNSOM recovery of the object function in the presence of noise. Color bar of Figure 1d also applies to panels a and f.

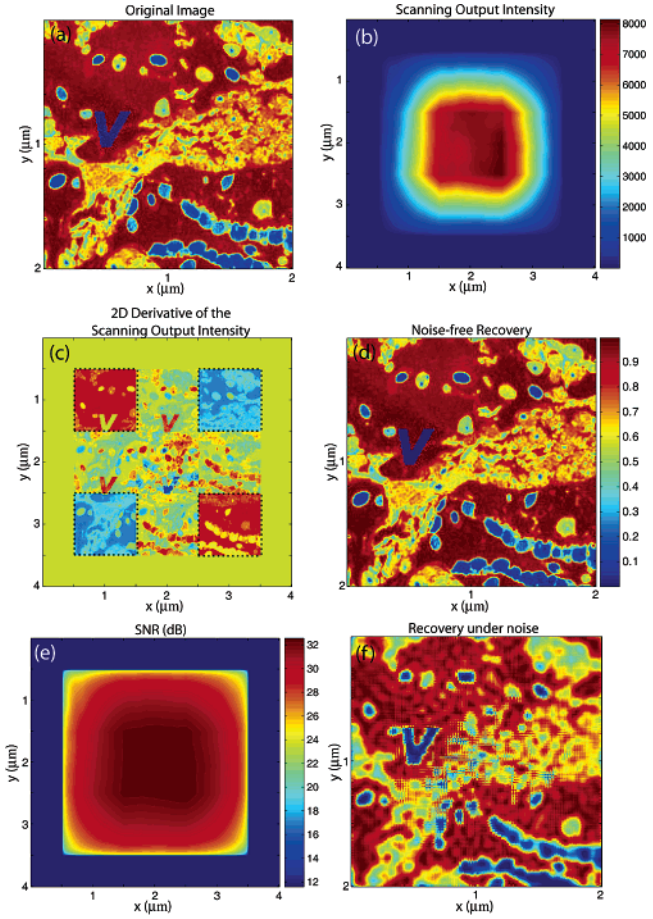

Figure 2. Same as Figure 1, except that this time the modeled DNSOM aperture was 1 μm × 1 μm. The original image area was 2 μm × 2 μm as before. The scanning output intensity together with its 2D derivative differed from that of parts b and c of Figure 1, respectively. Specifically, in Figure 2c, the center cross region could not be used due to significant spatial aliasing. Only the regions within the dotted gray squares contained unique spatial information about the original image. The color bar of panel d also applies to panels a and f.

right at the surface of the object and the near-field interaction between the object and the aperture was ignored. Since the aperture was the same size as the object, all the features of $O(x,y)$ were averaged out, as can be seen in Figure 1b. Notice also that the total extent of the scanning image is 4 μm × 4 μm, which is the required size for the 2 μm × 2 μm image to be fully outside the DNSOM aperture. Following eq 2, a 2D numerical derivative of Figure 1b was computed. The result of this computation is illustrated in Figure 1c, where red and blue colors represent positive and negative values, respectively. As expected from eq 2, four replicas (two positive and two negative) of the original image appear centered around four corners of the 2 μm × 2 μm square aperture.

For this numerical example, the width of the square was equal to the width of the object function, $O(x,y)$. Therefore, there was no spatial aliasing[22] in Figure 1c among the four replicas of $O(x,y)$. The final recovery of $O(x,y)$ was obtained by averaging the four replicas shown in Figure 1c, the result of which is depicted in Figure 1d. The error in this recovery, defined as $\epsilon = \int\int |O_r(x,y) - O(x,y)|^2 \, dx \, dy / \int\int |O(x,y)|^2 \, dx \, dy$, where $O_r(x,y)$ represents the recovered quantity, was negligible ($\ll 10^{-20}$%).

To further illustrate the effect of the aperture width, W, on the recovery, we ran another simulation using the same object function, with a square aperture width of $W = 1$ μm. The scanning output for this case (Figure 2b) was different than that of the previous example (Figure 1b); however, it still did not show any of the fine details of $O(x,y)$. Taking a 2D derivative of Figure 2b yielded the profile shown in Figure 2c. In this image, the center cross region could not be used for the recovery of $O(x,y)$ due to spatial aliasing,[22] where all four replicas of $O(x,y)$ now overlap (Figure 2c).

However, there exist four distinct regions (shown *within* the dotted gray squares in Figure 2c) which contain unique information about $O(x,y)$. By tiling these four smaller squares within the dotted regions of Figure 2c, $O(x,y)$ can now be recovered without loss of information (Figure 2d). Once again, the recovery in Figure 2d is excellent, with a negligible error of $\epsilon \ll 10^{-20}\%$. We should emphasize here that for these two numerical simulations reported in Figures 1 and 2, the ***only*** piece of information used to recover the original image was Figure 1b and Figure 2b, respectively. Since the central region in Figure 2c was not used in this recovery due to spatial overlapping, the measurement of the scanning power only needs to be done across regions enclosed within the dashed squares of Figure 2c, which can reduce the scanning time considerably.

Next, to simulate the noise performance of DNSOM, we assumed a shot-noise-like noise behavior where SNR at each scanning position was proportional to the square root of the total collected power. We limited the maximum SNR of all the pixels of the scanning output to $<\approx$35 dB, corresponding to an SNR that would be theoretically expected for an aperture size of $<$100 nm. This level of noise represents a worst-case scenario; we would expect our shot-noise-limited SNR for DNSOM aperture sizes $\geq$0.5 μm to be $>$25 dB higher. Under these assumptions, SNR maps, degraded by additive Gaussian noise for $W = 2$ μm and $W = 1$ μm are shown in Figures 1e and 2e, respectively. The corresponding recovery is illustrated in Figures 1f and 2f, for $W = 2$ μm and $W = 1$ μm, respectively. To reduce the noise sensitivity of $\partial^2/\partial x' \partial y'$ operation, a wavelet transform based filter[25] and a 2D Fourier domain filter were applied to the scanning output images. The second Fourier domain filter rejected most of the noise spatial frequencies that do not originally exist in $O(x,y)$, without affecting the accuracy of the $\partial^2/\partial x' \partial y'$ operation by passing all the high spatial frequencies along x and y. In addition to the filtering operations, the recovered image quality is also improved by combining the four replicas in the final recovery step, since each image has different noise terms. The recovery results depicted in Figures 1f and 2f show many details of the original image even in the presence of relatively large amounts of noise. The recovery errors were $\epsilon < 1.5\%$ and $\epsilon < 1.7\%$, respectively. Overall, our simulation results indicate that, under a fairly large amount of noise, the recovery of $O(x,y)$ is still quite good, suggesting the possibility of practical near-field imaging based on DNSOM.

Next, we experimentally demonstrated DNSOM. Ordinarily, we would scan DNSOM's square aperture (or detector), e.g., positioned on a cantilever tip, while the sample is illuminated in either transmission or reflection mode. However, owing to the simplicity of the experimental apparatus, we instead opted to demonstrate the proof-of-principle of DNSOM by imaging the near-field intensity distribution of a scanned, illuminated conventional NSOM tip (see Figure 3a). The NSOM tip was on a silicon cantilever with a hollow aluminum pyramid, with a hole diameter of $A = \approx$100 nm (Figure 3b). In our proof-of-concept experiment, a $\approx$1 μm wide DNSOM square aperture was fabricated

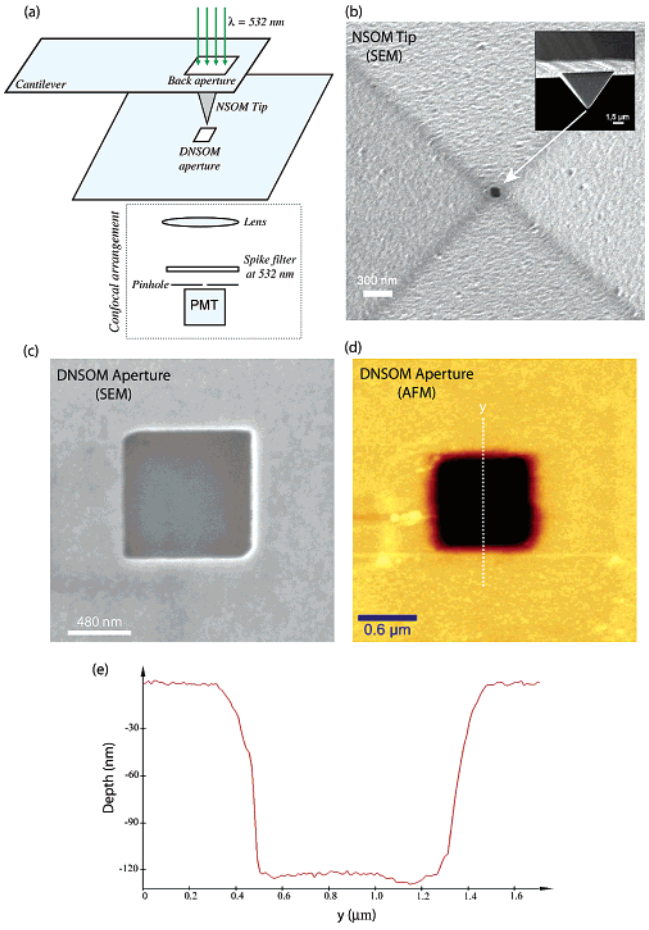

Figure 3. (a) Simplified schematic of the experimental setup. (b) SEM image of a typical NSOM tip used in this work. (c) SEM image of the fabricated DNSOM square aperture. (d) AFM image of the same fabricated DNSOM square aperture. (e) Cross section (along the white dotted line) of the AFM image shown in Figure 3d.

using focused ion beams onto a planar glass substrate, which was initially coated with a gold layer of $\approx$100 nm thickness. A scanning electron microscope (SEM) image together with an atomic force microscope (AFM) image of this fabricated aperture is shown in parts c and d of Figure 3, respectively. For the square aperture, we used $W = 1$ μm in order to ensure that the pyramidal geometry of the NSOM tip would be free from obstacles (e.g., square edges) while scanning the *floor* of the 100 nm thick aperture. Here, we should emphasize that the current experimental configuration of Figure 3a is chosen just to show the proof-of-principle of DNSOM and is *by no means* ideal. In the current planar aperture configuration (Figure 3a), the gold film was chosen just thick enough ($\approx$100 nm) to block light penetration outside of the aperture region. If the thickness of the gold film was made larger, this would result in increased optical losses within the aperture region, together with scanning problems due to increased height of the walls of the aperture, forming an obstacle for the continuous scanning of the NSOM tip. Meanwhile, this issue of film thickness becomes irrelevant in an ideal experimental configuration, where, e.g., DNSOM aperture is actually fabricated on the top of a conventional NSOM tip. Then, due to the pyramidal geometry of the

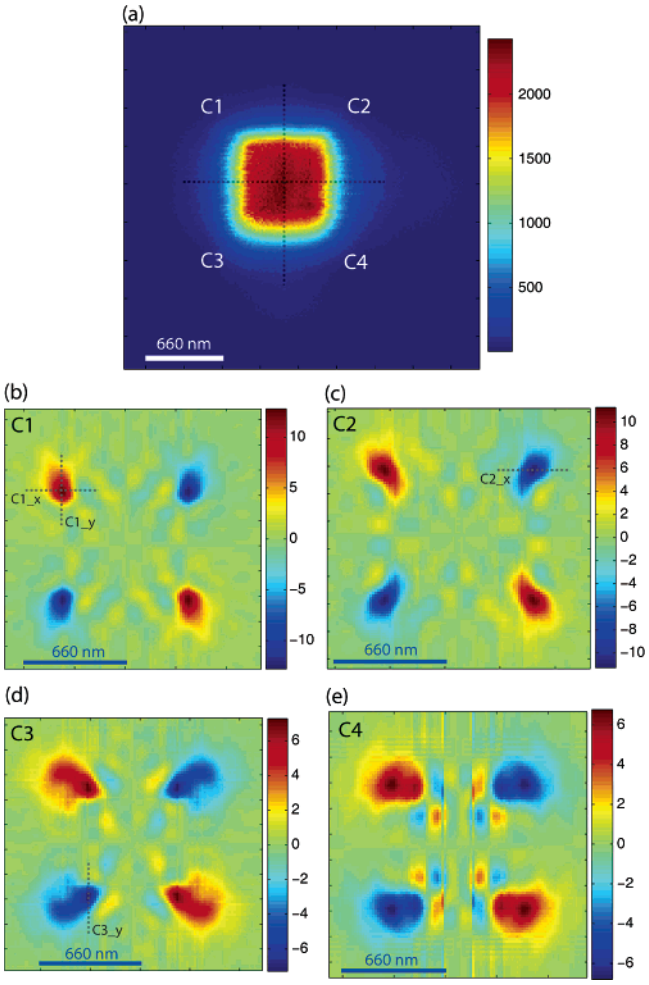

Figure 4. (a) Measured DNSOM image (scanning output power), where C1, C2, C3, and C4 denote the corners of the DNSOM square aperture. The data at each corner of this DNSOM image were analyzed separately. The results of this analysis are shown in panels b−e, corresponding to C1, C2, C3, and C4, respectively.

DNSOM tip, the above addressed problems will be partially eliminated.

In our measurements, a conventional aperture-type NSOM system (Witec, Alpha-SNOM) was used to scan the tip across the DNSOM aperture (Figure 3a). The angle of the cantilever with respect to the square aperture plane was ≈8°. Illumination of the NSOM tip was at 532 nm (linearly polarized before hitting the back aperture of the tip), and the tip was kept in the near-field of the square aperture using the ac feedback (tapping) mode. The scanning step size was ≈16 nm and the total range of scan was 4.2 μm × 4.2 μm. The transmitted light through the DNSOM aperture was collected by an inverted confocal microscope and detected by a photomultiplier tube (PMT). Figure 3a shows a simplified sketch of the experimental setup.

Figure 4a shows the DNSOM image that was obtained by measuring the power transmitted through the aperture as the NSOM tip was scanned over the DNSOM aperture. This power map has four corners, which we will denote as C1, C2, C3, and C4 (also illustrated in Figure 4a). Since for this experiment $W \gg A \approx 100$ nm, spatial aliasing between different corners, discussed above, was not present. Since each of these corners contained a unique and *independent* map of the intensity distribution of the NSOM tip's light transmission, we analyzed the data around each corner separately. Each quarter of Figure 4a, separated by the dashed black lines, was isolated and each of these quarter images (for instance C1) was individually flipped up/down and left/right to tile a new DNSOM image, which looked quite similar to Figure 4a, except that this time all corners were the same. In this manner, four different DNSOM images were formed, each one corresponding to only one corner information of C1, C2, C3, or C4. These four different DNSOM images (not shown) form the basis (before 2D differentiation) to parts b−e of Figure 4, respectively. Since $W \gg A$, the exact choice of the location of the borders between different corners (i.e., the dashed lines in Figure 4a) was not critical. Separate analysis of each corner of Figure 4a was useful for several reasons: (1) the effect of the fabrication imperfections such as angled side walls of the square aperture was reduced; (2) the effect of uncontrolled mechanical drifts during the scanning procedure was minimized; (3) each corner could be filtered separately with different, optimal digital filter parameters; and (4) information from different corners can be used to obtain a final image with improved SNR compared with that obtained by a single corner alone.

The 2D derivatives of each of the four processed (corner) images were computed. The recovery results for each corner (C1, C2, C3, and C4) are shown in Figure 4b−e. The same digital filtering discussed in the numerical simulation results was also applied here. In each of the recovery results of parts b−e of Figure 4, the profiles at the corners were spatially symmetric with respect to other corners, which is a natural result of the tiling operation discussed above. The recovery results of Figure 4 showed that the image quality obtained from corners C1 and C2 was better than that from C3 and C4, which can be attributed to either a structural change in the tip geometry or a potential problem with the ac feedback loop as the scanning process evolved (e.g., C1 was the first corner to be scanned whereas C4 was the last one). This degradation in C3 and C4 can also be visually seen in Figure 4a, where the corners C3 and C4 were more rounded and less well defined when compared with C1 or C2.

Figure 5 shows the *x* and *y* cross-sectional profiles of the images shown in parts b−d of Figure 4. The gray dotted lines in parts b, c, and d of Figure 4, refer to the locations where the cross sections were taken. Oscillatory artifacts in these profiles (see, e.g., Figure 5a) appeared when the NSOM tip was at the boundary between the metal layer and the aperture region. For instance, for C1, the oscillatory artifact appears at the right and for C2 it appears at the left of the peak as illustrated in Figure 5a. The fact that these artificial affects appear on different sides for C1 and C2 is beneficial; by taking the average of these two profiles, the effect of such ringing artifacts can be minimized. The result of this averaging for the C1_x and C2_x (see Figure 5a) is shown in Figure 5c with the solid blue line. The same phenomenon is also visible in the *y* direction between C1 and C3 as shown in Figure 5b. The averaged profile along *y* is also shown in Figure 5c with the dotted red curve. The full width at half-

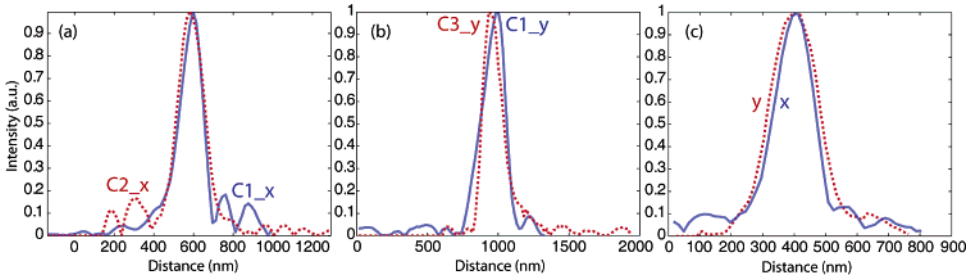

Figure 5. Cross-sectional profiles of the images shown in Figure 4. The gray dotted lines in panels b−d of Figure 4 refer to the location where the cross section is taken. (a) The *x* cross sections of C1 and C2 are shown. The result of averaging for C1_x and C2_x is shown in part c with the blue line. (b) The *y* cross sections of C1 and C3 are shown. The averaged profile along *y* is also shown in part c with the dotted red curve. The fwhm's along *x* and *y* were measured from part c and found to be ≈145 and ≈179 nm, respectively.

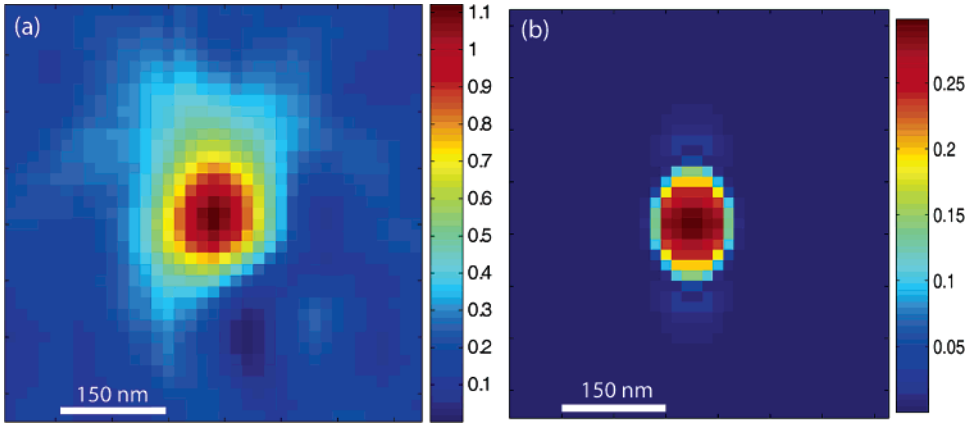

Figure 6. (a) Final recovery of the near-field intensity distribution of the 100 nm diameter NSOM tip (using C1 and C2). Each pixel corresponds to ≈16 nm. (b) FDTD simulation of the near field distribution of a 100 nm aperture on a planar aluminum film. The thickness of the aluminum layer was modeled to be 150 nm, and the illumination light was a plane wave with linear polarization at 532 nm. The measurement results are in fairly good agreement with the FDTD simulation results. The intensity units of the color bars are arbitrary.

maximum (fwhm) in Figure 5c along *x* and *y* was measured to be ≈145 and ≈179 nm, respectively. From this figure, since $A = 100$ nm, we can qualitatively argue that our approximate spatial resolution, *R*, in this experiment is $R \sim 50$ nm.

For the recovery of the final 2D intensity image, due to their superior performance we used the results of C1 and C2. By taking the average of the recovered images of C1 and C2 (parts b and c of Figure 4, respectively), we obtained the final 2D intensity map of the 100 nm NSOM tip as shown in Figure 6a. Overall, the recovery result agrees well with the physical size of the tip aperture (≈100 nm), where, including the skin depth (∼10 nm), the effective aperture diameter becomes ≈120 nm. Furthermore, the results are also in fairly good agreement with a finite difference time domain (FDTD) simulation of the near field distribution of a 100 nm hole opened on a 150 nm thick aluminum plane (Figure 6b). For this FDTD simulation, the illumination light is assumed to be a plane wave with linear polarization at 532 nm. We did not observe a major asymmetry between the *x* and *y* directions in our experimental results (Figure 6a), possibly because the polarization of the beam *exiting* the NSOM tip is not well defined in our experiment due to the pyramidical geometry of the tip.

It is quite interesting to note that the principles of DNSOM can also be extended to other aperture/detector geometries such as triangles and parallelograms. Especially triangles form an interesting group of aperture geometries that can be used to reduce the overall contribution of shot noise in the recovery result. For a triangular geometry, different than a rectangular (or square) structure, in an optimum DNSOM recovery process only a single corner will contribute to the retrieval of the 2D object function of interest. This corner of the triangle will have a height of *H* from its base and a full angle of α. In order to briefly outline the recovery process of a triangle aperture/detector based DNSOM system, let us denote the maximum width of the 2D object of interest as W_o. In this case, the optimum triangle geometry can be shown to have a height of $H = W_o$. The optimum corner angle, α, depends on the fabrication tolerances, the light throughput, and the skin depth requirements. The smaller α gets, the smaller the light throughput through the aperture becomes, and the harder it gets to fabricate a sharp distinct triangle corner. Furthermore, as α is reduced below 90° the two sides of the triangle corner start to *see* each other through electromagnetic coupling, making the skin depth effect stronger, as a result of which the effective resolution of the DNSOM recovery should decrease. In order to avoid such a penalty, $\alpha \geq 90°$ can be chosen. The numerical recovery process in a triangle-based DNSOM system will also need to be modified; i.e., the first spatial derivative taken along a direction that is parallel to one side of the top corner of the triangle will have to be followed by an angular rotation of the image, by α degrees, before a second derivative along

the same original direction is taken. If α is chosen to be 90°, then the same numerical recipe as in a square geometry based DNSOM can be applied without an image rotation. This way the top corner of the triangle ideally produces a Dirac-delta function, $\delta(x,y)$, which is convolved with the 2D object function of interest, providing the image recovery. The other two corners of the triangle do not cause any spatial information loss since $H = W_o$, which avoids spatial aliasing.

In conclusion, we can summarize important advantages of DNSOM as follows: (1) The resolution of the image is not limited by the size of the aperture, but rather it is practically limited by the sharpness of the corners of the square. This feature of DNSOM may be particularly relevant because it is easier to fabricate sharp corners in a relatively large aperture (or detector) than to fabricate small area apertures/detectors (with $A < 100$ nm). (2) DNSOM opens up the possibility of scanning a relatively large square-area detector (or an array of detectors) in the near field of an object to achieve nanometer resolution imaging. This characteristic of DNSOM may permit the detection of the near-field intensities in the near field, avoiding complexities and spatial frequency information mixing that occur when transferring the near-field information to far-field detectors. (3) The limiting effect of the skin depth on resolution is improved; i.e., in a small area aperture, the lower bound of the effective aperture width is roughly $2L_s$, whereas in DNSOM, the same limiting effect of the skin depth is reduced to $\sim L_s$. (4) The light throughput is increased since DNSOM does not require an aperture diameter of less than 100 nm to achieve nanometer-level resolution. (5) Low power damage threshold of the conventional small area NSOM probes should be improved with the larger area of a DNSOM aperture.

While the advantages of DNSOM are substantial, some limitations of this technique exist. First, the object size cannot extend beyond twice the aperture/detector width. This limitation does not exist in conventional NSOM. However, for an aperture size of, e.g., 1 μm × 1 μm, a corresponding object area of 2 μm × 2 μm would be adequate for many nanoimaging applications. An alternative solution for flat samples can also be use of a movable rectangular mask that has four times the area of a DNSOM aperture/detector. This way, larger area flat objects can be fully scanned using a DNSOM aperture by the help of the moving mask. In a similar manner, an array of apertures or detectors could also be used to extend the object area that can be scanned with DNSOM.

A second issue with DNSOM is that stabilization of the scanning position within the near field of the object of interest can be difficult for larger aperture/detector sizes. With conventional NSOM this issue of stabilization is less problematic since the tip size is $<\approx$100 nm. Furthermore, if the angle of the DNSOM aperture/detector plane is not well controlled with respect to the object surface, some corners of the square to may be away from the near field of the object, causing resolution loss. However, there exists still one corner of the square aperture/detector that will be kept closest to the object's near field, enabling high-resolution imaging from that corner. Full-resolution from all four corners can be achieved by improving the mechanical stability and precision of the scanning stage. And a final issue with DNSOM is the selection of the optimal aperture size for a given object, which is strongly dependent on the object size and SNR of the detection system. For small objects (with a size of $\sim\lambda$), increasing the DNSOM aperture width beyond λ is essentially unnecessary. The upper bound of the DNSOM aperture width is governed by the dynamic range, sensitivity, and the analog-to-digital conversion step of the detection system.

In conclusion, we have numerically and experimentally illustrated a new form of aperture-type NSOM. DNSOM involves using a relatively large size square (or rectangular in general) aperture/detector and recording the power collected from this large square structure as a function of the scanning position. Image recovery is achieved by taking a 2D derivative of the light detected through the scanning aperture/detector. A proof-of-principle of DNSOM was provided by imaging the 2D intensity distribution of $a \approx$100 nm diameter aperture NSOM tip using a square DNSOM aperture of $\approx$1 μm width. Our results revealed a fwhm of $\approx$150 nm for the near-field distribution of the tip. Unlike conventional aperture-type NSOM, the size of the rectangular aperture/detector does not determine the resolution of the recovered image. In DNSOM, the resolution is determined by the sharpness of the corners of the rectangular aperture/detector. These features of DNSOM make it potentially advantageous for nanometer-level imaging, especially when resolution and light throughput are at a premium.

Acknowledgment. We acknowledge M. Kress of Witec for his excellent technical support on NSOM imaging. Two of the authors (E.C. and F.C.) are supported by the AFOSR under a MURI program. The support by the NSF NSEC and the Center for Nanoscale Systems (CNS) at Harvard University is also gratefully acknowledged.

References

(1) Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press: New York, 2006.
(2) Lewis, A.; Isaacson, M.; Harootunian, A.; Muray, A. *Ultramicroscopy* **1984**, 13, 227.
(3) Pohl, D. W.; Denk, W.; Lanz, M. *Appl. Phys. Lett.* **1984**, *44*, 651.
(4) Fischer, U. Ch. *J. Vac. Sci. Technol., B* **1985**, *3*, 386.
(5) Betzig, E.; Lewis, A.; Harootunian, A.; Isaacson, M.; Kratschmer, E. *Biophys. J.* **1986**, *49*, 269.
(6) Betzig, E.; Trautman, J. K.; Harris, T. D.; Weiner, J. S.; Kostelak, R. L. *Science* **1991**, *251*, 1468.
(7) Betzig, E.; Chichester, R. J. *Science* **1993**, *262*, 1422.
(8) Betzig, E.; Finn, P. L.; Weiner, J. S. *Appl. Phys. Lett.* **1992**, *60*, 2484.
(9) Inouye, Y.; Kawata, S. *Opt. Lett.* **1994**, *19*, 159.
(10) Ephrat, P.; Roodenko, K.; Nagli, L.; Katzir, A. *Appl. Phys. Lett.* **2004**, *84*, 637.
(11) Rasmussen A.; Deckert V. *Anal. Bioanal. Chem.* **2005**, *381*, 165.
(12) Mitsui, T. *Rev. Sci. Instrum.* **2005**, *76*, 043703−1.
(13) Gan, Q.; et al. *Appl. Phys. Lett.* **2006**, *88*, 121111.
(14) Bethe, H. A. *Phys. Rev.* **1944**, *66*, 163.
(15) Garcia de Abajo, F. J. *Opt. Express* **2002**, *10*, 1475.
(16) Heng, X., Erickson, D.; Demetri, P.; Changhuei, Y. *Proc. SPIE.* **2005**, *6003*, 40.
(17) Zenhausern, F.; Martin, Y.; Wickramasinghe, H. K. *Science* **1995**, *269*, 1083.
(18) Knoll, B.; Keilmann, F. *Nature* **1999**, *399*, 134.
(19) Yang, T. J.; Lessard, G. A.; Quake, S. R. *Appl. Phys. Lett.* **2000**, *76*, 378.

(20) Keilmann, F.; Hillenbrand, R. *Philos. Trans. R. Soc. London, Ser. A* **2004**, *362*, 787.
(21) Barbara, A.; Lopez-Rios, T.; Quemerais, P. *Rev. Sci. Instrum.* **2005**, *76*, 023704−1.
(22) Goodman, J. W. *Introduction to Fourier Optics*; McGraw-Hill: New York, 1996.
(23) Carney, P. S.; Schotland, J. C. *Appl. Phys. Lett.* **2000**, *77*, 2798.
(24) Carney, P. S.; Markel, A. M.; Schotland, J. C. *Phys. Rev. Lett.* **2001**, *86*, 5874.
(25) Starck, J. L.; Murtagh, F.; Bijaoui, A. *Image Processing and Data Analysis: The Multiscale Approach*; Cambridge University Press: Cambridge, U.K., 1998.

NL062110V