Statistical approach for subwavelength measurements with a conventional light microscope

Daniel Palanker and Aaron Lewis
Division of Applied Physics, The Hebrew University of Jerusalem, Jerusalem, Israel

ABSTRACT A method is developed theoretically that will permit subwavelength measurements of objects that differ from the surroundings by any contrast enhancing parameter, such as fluorescence, second harmonic generation, reflection et cetera, using a statistical analysis of a picture obtained with a conventional light microscope through a set of subwavelength apertures or by repeated scanning of a laser beam over a defined area. It is demonstrated that with this methodology information can be obtained on microdomains that are thirty times less than the diameter of the aperture. For example, for apertures that are 0.3 μm in diameter it is possible to measure the dimension of objects that are ~10 nm. A technology is described by which it is possible to produce masks with the appropriate apertures. Instrumentation is described that would allow for the realization of these statistical methodologies with either apertures or scanning laser beams.

INTRODUCTION

There has been a considerable effort expended on obtaining higher resolution information with conventional light microscopes. These efforts can be classified into three general categories which are interference techniques (1), confocal methodologies (2–4) and computational approaches with charge coupled devices as detectors (5–7). Such efforts have resulted in considerable improvements in resolving biological structures that are at the very limit of light microscopy. This limit can generally be defined as 0.25 μm for visible light. In spite of these advances there are numerous structures and aggregates of molecules that cannot be resolved with methodologies that rely on lens based light microscopes that are limited in their ultimate resolution by the wavelength of light. Examples of these structures include microdomains in membranes, distribution of fluorescent properties in chloroplasts and mitochondria, ribosomes, nuclear knobs, et cetera that can only be studied in the electron microscope.

Recently, a new lensless method of light microscopy has been developed that has the potential of obtaining higher resolution than the diffraction limit associated with lens-based instruments. This type of microscopy has been called near-field scanning optical microscopy (NSOM) (5, 8–14). In NSOM, an aperture is scanned in close proximity to the surface that is being imaged. The resolution of the technique depends on the aperture diameter and its distance from the surface. For resolutions that are significantly better than the limits of conventional light microscopes aperture/surface separations of a few tens of nanometers have to be maintained and this, in fact, is the principal difficulty of NSOM.

In this paper we introduce a new method which is not based on scanning a single aperture but rather relies on a statistical analysis of information obtained from an assembly of similar objects viewed in a standard light microscope through a mask of submicron apertures. It is shown that such an analysis allows for the measurement of the dimensionality of objects with an accuracy that is dependent on the number of apertures within the field of view. For example, analyzing a picture with 100,000 apertures makes it possible to measure the size of objects with an accuracy of ~0.8% of the diameter of the holes in the mask. Alternately, we show that similar information may be obtained by applying analogous approach to data obtained by repeatedly scanning a laser beam over a field of view. The methods we have developed have the potential to determine diameters on subwavelength molecular aggregates or structures in a hydrated form at resolutions that have generally been limited to electron microscopy of dehydrated materials. In addition, when aperture arrays are used this information on aggregate dimensionalities is obtained without interference from out-of-focus contributions from below the surface that is being interrogated. This near-field effect is due to an exponential decrease in light intensity with distance from the aperture (15) and as a result only objects in close proximity to the apertures are detected (10). As mentioned above the near-field region is approximately equal to the diameter of the holes in the screen. To describe this new statistical approach to superresolution measurements of dimensionalities with a conventional light microscope this paper will be divided into
two parts that deal with two distinct methods of data collection and analysis. Finally, experimental methods will be described that will allow for the application of this statistical approach to light microscopy.

**METHODOLOGY**

**Statistical microscopy with apertures**

**A mathematical technique**

Consider a plane containing many objects that differ from their surroundings by some contrast-enhancing parameter such as fluorescence, second harmonic generation et cetera. The dimension of these objects is smaller than the resolving power of lens-based microscopes. In the discussion below let us consider, for example, fluorescence. The plane of these fluorescent entities is covered with an opaque, thin, flat screen which contains uniform holes with a diameter $D$ (See Fig. 1A). Above this plate with holes is placed the objective lens of a fluorescence microscope. Only those holes that cover the fluorescent objects or parts of them are visible as dots of light by the detector of the microscope (see Fig. 1B). Because the holes in the opaque screen could have dimensions that are less than the resolving power of the microscope they are detected as point sources of light that are diffraction limited. To be detected through a hole the object must be present, at least partially, under the aperture. To derive the equations that could give the probability of this event, a model is built in which an arbitrary point on an object is chosen as a "fixed point." The next step is to move the model object around the border of the hole while maintaining its orientation in the plane (see Fig. 2A). In this operation the fixed point draws out the border of some figure (dotted line in Fig. 2A) with a surface area that is labeled $\sigma$. For convenience, let us call the figure represented by the dotted line in Fig. 2A "F." In the case of another orientation of the object, the same figure will be obtained except for an appropriate change in orientation. This arises due to the round shape of the hole. Thus, the criterion for detection of a randomly located object through the hole is that its fixed point will fall within the borders of figure $F$ oriented in a way dictated by the object orientation. The probability that none of the randomly distributed objects will fulfill this criterion and thus will not illuminate the hole is given by a Poisson distribution which is $\phi = e^{-\sigma}$, where $\epsilon$ is the lateral concentration of the objects on the plane. It is possible to use this approximation under the condition that the objects are indeed randomly distributed and do not effect one another. Simulations on a computer show that

**FIGURE 1** (A) A plane wave of exciting radiation is transmitted through a subwavelength hole in an opaque screen onto a disc which represents a collection of molecules differing from their surroundings by some contrast enhancing parameter such as fluorescence. The molecules emit light through the same hole and this light is collected by the lens of a conventional fluorescent microscope. (B) If the subwavelength aperture covers any part of the fluorescent disc, then the detector sees it as a dot of light with an intensity that is dependent on the area of the disc that is within the aperture diameter.

![Diagram](image1)

**FIGURE 2** (A) A point in the object called a "fixed point" draws out a border for a characteristic figure $F$ (dotted line) when the object moves around the subwavelength aperture touching the border of the aperture and preserving a constant orientation in a plane. $\sigma$ is a surface area of the figure $F$. (B) Figure F in the special case of a disc-shaped object is a circle with a surface area $\sigma = \pi (D + d)^2/4$
this approximation works well for coverages of the plane by the objects that are <0.1 of the area. The ratio of the apertures that are not lighted by the fluorescent entities (labeled \( N_{\text{tot}} \)) to the total number of the holes, \( N_{\text{tot}} \), is given by

\[
N_{\text{tot}}/N_{\text{tot}} = e^{-\sigma} \quad (1)
\]

By measuring \( N_{\text{tot}} \) and \( N_{\text{tot}} \), \( \sigma \) can be calculated and by knowing \( D \) it is possible to calculate the dimension of the objects. For this calculation it is required to obtain a value for \( \epsilon \) and to know the relation between \( \sigma \) and the object’s dimensions. If the object’s shape is known apriori, this relation is possible to evaluate exactly. If not, we can estimate its dimension in the following way. In place of the object we can take a disc with the “fixed point” at its center that draws around the aperture a figure with the same surface area \( \sigma \) as above (see Fig. 2 B). The diameter \( d \) of this disc is a good approximation of the average object’s dimension. For example, if the objects are of elliptical shape with axes \( a \) and \( b \), \( d \) will be a rather good approximation of the average of these two lengths (i.e., \( (a + b)/2 \sim d \)). The error is equivalent to 0 in the case of \( a/b = 1 \) and is \( \sim 30\% \) in the other extreme of the ellipse where \( a/b = 0 \). In the case of the disc-shaped object the relation between \( \sigma \) and \( d \) is very simple:

\[
\sigma = \pi(D + d)^2/4. \quad (2)
\]

All of the subsequent calculations in this article will assume disc-shaped objects in view of the fact that we have already explained the relationship between this specific case and objects with other shapes.

For obtaining \( \epsilon \) there are two possibilities. The first is to directly count the quantity of the fluorescent discs on some area of the plane without the screen even if the exact diameter is not resolvable. This directly gives \( d \) by Eqs. 1 and 2. The statistical character of the number of lighted apertures naturally causes a statistical error in the calculation of \( d \) with a dispersion

\[
\delta = \Delta \sigma/\sigma = e^{x}\sqrt{1 - e^{-\sigma}}/(a\sqrt{N_{\text{tot}}}) \quad (3)
\]

where

\[
a = \epsilon \sigma. \quad (4)
\]

The statistical error for the measurement of \( d \) is

\[
\Delta d/(d + D) = \frac{1}{2}\delta. \quad (5)
\]

The problem with this method is the fact that when the number of discs is counted without the screen it will be greater than with the screen, not only because of the lack of coverage, but also because of the near-field effect of the aperture. This occurs because the electromagnetic field experiences an exponential decrease as a function of distance from the aperture plane (10, 15). Thus, it is possible to use the method of direct counting of the fluorescent disc number only for the analysis of objects that are distributed in an \( xy \) plane but not in \( z \). An example of a fluorescent object that would fit this criteria is a microdomain in a monolayer or bilayer membrane.

Because of the above limitation, a method has been developed to bypass the need to directly measure \( \epsilon \). In this method, that we call the three screens method, the ratio \( f = \ln (N_{\text{tot}}/N_{\text{tot}}) = \sigma \) is measured for three screens each with uniform holes of diameters \( D_1, D_2 \) and \( D_3 \). From these measurements we can calculate a derivative \( f' = df/\delta D \) at a midpoint \( D_2 \) and to derive \( d \) by the following equations:

\[
d = 2(f_1/f_1') - D_2, \quad (6)
\]

where

\[
f_1' = [(f_1 - f_2)(D_2 - D_1)^2 + (f_2 - f_3)(D_3 - D_1)^2]/
[(D_3 - D_1)(D_2 - D_1)(D_3 - D_1)], \quad (7)
\]

with a statistical error of

\[
\Delta d/(d + D_1) = \delta[1 + A(D_2 + d)^2/4 - B(D_2 + d)]^{1/2}, \quad (8)
\]

where

\[
A = [(D_2 - D_1)^2 + [(D_1 - D_2)^2 - (D_2 - D_1)^2]]/(D_3 - D_1)^2,
\]

and

\[
B = [(D_1 - D_2)^2 - (D_2 - D_1)^2]/
[(D_1 - D_2)(D_2 - D_1)(D_3 - D_1)]. \quad (9)
\]

Statistical microscopy does not require one to resolve the dimension of the hole through which the fluorescence is being observed. Rather, it requires only the recognition of the holes as points of light. Therefore, it is possible to use a low power microscope lens such as 10× or even lower objective for holes of diameter <1 \( \mu \)m and such low magnification provides a large field for statistical analysis. For a microscope with an objective lens that has a magnification of 10× it is possible to observe a region of 1,000 \( \mu \)m \( \times \) 1,000 \( \mu \)m and this area may contain \( N_{\text{tot}} = 10^5 \) of the holes with a concentration of 0.1 \( \mu \)m\(^2\). This large area yields a substantial data base \( N_{\text{tot}} \) that provides high accuracy.

For \( N_{\text{tot}} = 10^5 \), we calculated the accuracy of statistical microscopy (SM) with formulas 6–10 of the second method above which we call the three screens method. The results are shown in Fig. 3 where the curve is
derived from the formulas above and represents the dependence of the accuracy of the measurement of the diameter of the fluorescent discs on the lateral concentration of the discs. An entire series of experiments was simulated where a random distribution of holes and discs was provided by the computer and the results obtained are represented by asterisks in Fig. 3. Even though the theoretical calculation is for 100,000 holes, we completed the simulation for only 4,000 holes and we then scaled the results to 100,000 holes. This was done to save computer memory and time. As can be seen the simulation gave results that are very close to the theoretical curve plotted. These results were acquired for the region where the highest accuracy could be obtained and thus it is encouraging that the simulation provides good support for the theoretical framework we have developed. Our results show that it is possible to measure a diameter of the discs $d$ less than $0.5D_z$ with absolute accuracy $\Delta d = 0.008D_z$. This says, for example, that for $d = 0.1D_z$ the relative accuracy will be $\Delta d/d = 0.08$. Alternately, for $d = 0.03D_z$ we obtain $\Delta d/d = 0.25$. In other words, by putting the screen with the holes on a surface with fluorescent discs that are not resolvable with any presently employed methods of fluorescence microscopy we can see (i.e., measure the diameter) with a resolution of $30x$ better than the diameter of the holes in the plate. For example, using a set of three plates with apertures that are $0.2 \mu m$ in one plate, $0.3 \mu m$ in the second plate, and $0.45 \mu m$ in the third, it is possible to measure the dimension of fluorescent objects that have diameters of $10 \text{nm}$ with an accuracy of $2.5 \text{nm}$.

As is seen in Fig. 3, the accuracy with respect to the density of the objects is limited on the left side of the curve because of the small number of objects in the field of view and on the right side because at these concentrations nearly all the holes are lighted irrespective of the diameter of the fluorescent discs and because of interactions between the discs on the surface that cause a nonrandomness in the distribution. Our computer simulation also shows that in statistical microscopy the average diameter of the discs is measured and this does not depend on the dispersion of the diameter around the average.

It is understandable that if the fluorescent objects are distributed randomly on the plane, the randomness of the distribution of the holes in the screen have no effect. However, if the objects are placed in some particular order it is necessary to use screens with randomly distributed holes.

The statistical analysis requires the assignment of all apertures below a certain light intensity to be considered black and those apertures with intensities above this threshold to be considered illuminated. This can be accomplished by introducing a level of discrimination of the signal by the detector. Such a discrimination has the effect of reducing, by a small amount, the diameter of the hole and will cause an error in the object diameter measurements. This can be corrected by choosing different discriminator levels and in an iterative way the error induced by this choice can be elucidated. The same effect in the reduction of the aperture diameter can be introduced when the metal aperture border quenches the fluorescence from molecules that are in close proximity with the border.

**A histogram method**

An alternate possibility for obtaining information in a qualitative sense on the diameter of the discs is the use of histograms to analyze the intensity of the light emerging from the illuminated holes. In the computer simulation of this form of SM, we assume the same brightness (i.e., intensity per unit area) of the objects and a linear relationship of the intensity of the light emitted from the hole with the area of the fluorescent disc that is covered by the hole (see Fig. 1). This linear relationship is operative only for object/aperture separations corresponding to the near-field and when the possibility of quenching of the fluorescence by the metal of the screen is not considered. A more accurate
definition of the relationship between the light intensity and the area of the disc that is observed through the aperture, which takes into account finite distances between the screen and the object and the quenching effect, requires more detailed calculations (16).

In Fig. 4, A–F, is seen the qualitative nature of the histogram method. In this Figure is plotted the simulated intensities for different diameters $d$ of the discs on the surface. We can see that the relationship of the heights of the peaks with different intensities of light is dependent on the diameter of the discs. The presence of the peaks in the middle of the histogram is characteristic of a condition where the aperture diameter, $D$, is greater than the diameter of the discs, $d$ (see Fig. 4 A–C). If the histogram decreases monotonically from the origin, it shows that $d = D$ (see Fig. 4 D). If a peak occurs at

![Diagram](image)

**Figure 4** Simulation of the histogram method for different diameters of the discs. The abscissa of the graph represents a light intensity measured in 64 levels and the ordinate represents the number of illuminated holes. In A–C, the maximum intensity (channel number 64) is the intensity of three discs whereas in D–F, the maximum intensity is the signal detected when the hole fully covers a disc. In each case, the number of holes used was 4,000 and the relationship of the diameters of the discs to that of the holes in A–F was, respectively, 0.1, 0.25, 0.5, 1, 2, and 4.
the end of the histogram it indicates that $d > D$ (see Fig. 4 E–F). As can be seen in Fig. 5, a serious problem with the histogram method is the fact that as the dispersion in the diameters of the discs grows the peak of the histogram becomes both less intense and wider. This means that the histogram method can only be used to determine accurately the diameter of the discs if they are very uniform (i.e., differences of <10% in their diameters). It should be noted, however, that if the sample is composed of particles of two different sizes we should see two sets of peaks corresponding to each particle type.

### Scanning statistical microscopy

In addition to the above forms of statistical microscopy there is the possibility of a method of scanning statistical microscopy (SSM). Assume that a laser beam of diameter $D$ scans a surface with discs of diameter $d$ which differ from the surrounding by some contrast enhancing parameter. Because this technique requires repeated scanning of a sample by a laser beam, let us consider in the discussion below a parameter such as second harmonic generation that can be measured without bleaching the sample (see discussion in the last section). If a uniform step in this scanning procedure is made with accuracy $\Delta x > d$, it would be impossible to measure the diameter $d$. In SSM, a laser beam would transiently illuminate a specified position on a sample with defined discs as in Fig. 1 A with a spot of a laser beam in place of the hole. For the statistics in this method it is not necessary to scan over a wide field. Instead, we depend upon the inaccuracy of the stage to provide the statistical base on which our analysis will be built. In essence, in this method the stage scans with some inherent inaccuracy and thus the laser illuminates the sample with some statistical variation that generates the data base for our analysis. Thus, it is not necessary to scan a large area but rather it is necessary to have a large statistical assembly of the random positions of the discs and the beam. It is even possible to work with one such disc and the statistical assembly of the intensity is obtained by scanning over the selected region with a partial or random correlation between the positions of the laser beam and the disc. For this methodology, it is necessary to know the dimensionality of the surface area scanned.

If the quantity of the discs and the surface area scanned are known we can use the method described in the previous section with Eqs. 3–5. If these quantities are not known then the three screens method discussed above can be applied with three different spot sizes of the laser beam being used in place of the three screens with different hole sizes. Let us suppose that the distribution of the energy in the beam is Gaussian, $I(r) = \exp(-r^2/R^2)$, where $R$ is the radius of the beam, then the intensity of the light that illuminates an entire disc of diameter, $d$, that is placed at a distance, $l$, from the center of the beam is given by

$$I(l) \propto 2 \int_{-d}^{d} e^{-r^2/R^2} \cdot r \cdot \arccos \left( \frac{r^2 + d^2 - l^2}{2lr} \right) \, dr \quad l > d, \quad (11)$$
and

\[ I(l) \propto R^2 \left( 1 - e^{-i \theta - \eta l n^2} \right) \]

\[ + 2 \int_{1-R}^{1-R/3} e^{-i \theta \eta n^2 \cdot r \cdot r \cdot \arccos \frac{f^2 + r^2 - d^2}{2fr} dr } \quad l < d. \]  

(11')

In the event where \( d \ll R \) from Eq. 11, we get

\[ I(l) = \pi d^2 \exp \left( -\frac{f}{R^2} \right) \]  

(12)

In Eq. 12, we choose \( l = R \) for the direct counting method or \( l_1 = R_1, l_2 = R_2 \) and \( l_3 = R_3 \) for the three screens method. From the experimental data we choose a set of points with an intensity \( I \) such that \( I/I_{\text{max}} \leq \exp ( -1 ) = \alpha_0 \) and this is called \( N_{\text{bk}} \). Now from \( f = \ln (N_{\text{tot}} / N_{\text{bk}}) \) with the help of the direct counting or three screens method it is possible to obtain \( d_0 \) which is a first approximation of the diameter of the discs. Then \( d_0 \) is placed in Eqs. 11 and 11' and one gets from this \( I(l = R + d_0)/I_{\text{max}} = \alpha_0 \). By choosing a set of points from the experimental data which have intensities \( I \) such that \( I/I_{\text{max}} < \alpha_0 \), we obtain \( f = \ln (N_{\text{tot}} / N_{\text{bk}}) \) which in turn gives us \( d_1 \) as a second approximation of the diameter of the discs. The convergence of this process may depend upon the choice of \( \alpha_0 \). It may be that the laser beam is not precisely Gaussian. Then it is possible to try and apply SSM to known objects and to obtain a suitable function for \( I(r) \). It is possible to use the methodology developed here to not only investigate the diameter of discs but also to measure widths of lines that are defined by a contrast enhancing parameter such as second harmonic generation, reflection, etcetera.

In summary, statistical microscopy measures the statistical weight of the random superpositions of a laser beam or a hole with an object and this permits us to measure the diameter with a resolution that depends on the size of the ensemble. It is possible to obtain the same result as SSM if there is an ideal or very accurate (relative to the dimension of the object) xy stage in place of the randomness of practical stages and if the energy distribution in the beam is accurately known. With these parameters, it is possible to deconvolve the effect of beam shape from the distribution of the intensity obtained by scanning an object which is differentiated from its surroundings. Repetitive scanning in SSM overcomes the problem of the known inaccuracies that exist in the xy stage of scanning light microscopes. By appropriate manipulation of the mathematics above it is also possible to extend the theory to absorbing objects since all that is required is a difference in the interaction of light with the object and its surroundings.

**INSTRUMENTATION**

**Statistical microscopy with apertures**

An important aspect in the experimental realization of this statistical approach to subwavelength measurements is a technology to rapidly produce highly absorbing plates with uniform holes of dimensions smaller than the optical wavelength. In fact, this aspect of the experimental realization of statistical microscopy has already been accomplished. To obtain such plates, we have purchased latex spheres of uniform distribution from Sigma Chemical Co. (St. Louis, MO). These spheres were dispersed on a quartz slide (see Fig. 6A) and metalized (see Fig. 6B). Subsequently an excimer laser at 193 nm was transmitted through the quartz slide and was absorbed by the polymer molecules that compose the latex particles. This absorption caused a microexplosion in the vicinity of the latex sphere which ejects it with its metal cap (see Fig. 6C) leaving a transparent hole in the metal coating (see Fig. 6D).

Our method differs from an earlier approach used by Fischer (10) to produce submicron apertures in metal plates in which the latex spheres were removed from under the metal coating by sonication in a latex dissolving organic solvent. As noted by Fischer, this

![FIGURE 6](image-url)

A schematic representation of the method used to produce metal masks with uniform holes of submicron dimension. (A) Latent spheres with a diameter that is equal to the desired hole diameter are dispersed on a quartz slide. (B) Consequently a metal coating is evaporated on the side of the slide on which the spheres were deposited. (C) A pulse of 193 nm excimer laser radiation is transmitted through the slide and is absorbed by the latex causing microexplosions which eject the spheres. (D) Transparent holes of uniform diameter in the metal film on the quartz slide.
procedure is not an efficient method for the removal of subwavelength latex balls. For Fischer's purposes it was good enough because in his experiments he wanted to generate single submicron apertures and not the arrays that are required for statistical microscopy. Therefore, we developed the method described above which can readily produce the required screens. In Fig. 7 is shown such a mask of holes. Even though this plate does not have a uniform diameter of holes, the appropriately sized apertures can be chosen with either a CCD camera or a scanning laser microscope that is coupled to a computer with a program that is capable of selecting the holes on the plate with the correct dimensions. With such instrumentation, it is required to first choose the appropriate holes without a fluorescent filter in the microscope and subsequent to this choice to detect the fluorescence only through the chosen holes. This data is then analyzed by the methodologies described above. Hopefully we will be able to purchase in the future, the equipment necessary for this statistical analysis which will allow us to experimentally realize these ideas.

Scanning statistical microscopy

Presently, there are several commercial instruments that allow for a computer controlled laser beam to scan over specific regions of a surface and this enables the realization of SSM. In principle, unlike the case of statistical microscopy with apertures, there is no limitation for viewing only surfaces. One way to probe beneath the surface is to apply the advantages of confocal microscopy which allows the optical sectioning of the object being viewed (2–5, 7). However, if a small ensemble of fluorescent regions are used for the statistical analysis then there will undoubtedly be a significant problem of bleaching. It is possible that methodologies to prevent bleaching (17) could alleviate some of these problems. One way in which this serious limitation may be avoided would be to use a nonlinear spectroscopic phenomena such as second or third harmonic generation. These nonlinear methodologies can use resonances of the harmonic rather than the fundamental photon with absorptions in the sample. This would minimize bleaching while maintaining the characteristics of the point spread function that is characteristic of confocal techniques (2). In addition, as has been shown by Wilson and Shepard (2) harmonic microscopy has the ability to achieve z resolution that is comparable or better than confocal methodologies without the use of confocal apertures (2).

In summary, we have outlined in this paper a new approach to obtaining information on the dimension of subwavelength objects. The approach presented in this paper could be extended to investigate other parameters of importance in biological aggregates such as, for example, the fluidity of subwavelength membrane micro-
domains which could be probed by investigating through an aperture array the dynamics of recovery after photo-bleaching.

The authors would like to thank Professor Lazar Friedland for very helpful discussions with regard to a variety of mathematical techniques.

This research was supported by the United States Israel Binational Foundation and the German Israel Foundation.

Received for publication 26 November 1990 and in final form 22 April 1991.

REFERENCES