Fluorophore localization algorithms for super-resolution microscopy

Alex Small & Shane Stahlheber

Super-resolution localization microscopy methods provide powerful new capabilities for probing biology at the nanometer scale via fluorescence. These methods rely on two key innovations: switchable fluorophores (which blink on and off and can be sequentially imaged) and powerful localization algorithms (which estimate the positions of the fluorophores in the images). These techniques have spurred a flurry of innovation in algorithm development over the last several years. In this Review, we survey the fundamental issues for single-fluorophore fitting routines, localization algorithms based on principles other than fitting, three-dimensional imaging, dipole imaging and techniques for estimating fluorophore positions from images of multiple activated fluorophores. We offer practical advice for users and adopters of algorithms, and we identify areas for further development.

The resolution of a light microscope was once thought to be limited to about half the wavelength of light[1]: light waves from sources separated by less than that distance form overlapping blurs when imaged. The past several years have seen remarkable progress in overcoming this limit via several innovative super-resolution approaches, including localization of switchable fluorophores[2–5], localization of fluorophores from sequential images of bleachable fluorophores[6,7], saturation of upward transitions[8] and saturation of downward transitions[9,10]. Techniques based on localization of switchable fluorophores go by a number of names: for example, photoactivated localization microscopy (PALM)[3], stochastic optical reconstruction microscopy (STORM)[4] and fluorescence PALM (fPALM)[5]; we shall simply say 'localization microscopy'. Regardless of name, all of these approaches share two key innovations: (i) fluorophores that alternate stochastically between activated states (in which molecules fluoresce in response to excitation light) and dark states (in which molecules do not fluoresce) and (ii) algorithms that can localize individual fluorophores from images. A high-resolution image is constructed from the fluorophore positions; the resolution of the reconstruction, and its fidelity to the underlying specimen, depends on the density of the labels (the Nyquist criterion[11]) and the process of reconstructing the final image[12] as well as the precision with which the fluorophores are localized[12–14].

The density of the labeling and the precision with which fluorophores can be localized are both influenced by fluorophore properties, and the choice of fluorophore depends on the nature of the specimen and the target being labeled. Fluorophore development is therefore an area of broad and rapid innovation. Algorithms, on the other hand, tend to be applicable across broader categories of experiments and specimens. Localizing fluorescent probes predates super-resolution microscopy[15,16] and builds on prior work in astronomy and other fields[17,18]; some current algorithms[19] are even inspired by global positioning systems[20]. Not surprisingly, the advent of super-resolution microscopy, and the biological insights it enables, has led to a new wave of algorithm development. Issues important to localization microscopy include the stochastic on-off process[21], imaging in three dimensions (3D), the effects of dipole orientation[22] and the effects of local structure[23]. Localization microscopy is used for many structures, including fluorophores closely spaced along a line but sparsely spaced in the direction perpendicular to the line (such as microtubules[24] or even photolithographic structures[25]), point

Department of Physics and Astronomy, California State Polytechnic University, Pomona, California, USA. Correspondence should be addressed to A.S. (arsmall@csupomona.edu).

RECEIVED 22 MAY 2013; ACCEPTED 22 JANUARY 2014; PUBLISHED ONLINE 27 FEBRUARY 2014; CORRECTED AFTER PRINT 23 JULY 2014; DOI:10.1038/NMETH.2844

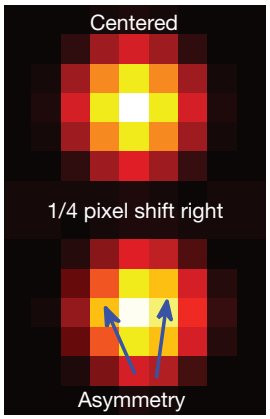

Figure 1 | Subpixel and subwavelength information: small shifts of the fluorophore alter the spatial distribution of light and the asymmetry of the image.

objects with correlated distributions (such as protein clustering[26]) or even clouds of diffusing ions (such as Ca^{2+} puffs[27]). Each situation poses unique challenges, and one should benchmark an imaging system and algorithm for the relevant type of structure[23]. Readers seeking inspiration are encouraged to examine the image processing literature of other fields, especially astronomy, as there are undoubtedly many gems not yet harnessed for super-resolution microscopy.

This Review begins with the concept of localizing a fluorophore, the effects of noise and background and the theoretical limit to localization precision. We then survey methods for localizing isotropic point sources in 2D; these methods are the most mature, and they illustrate issues common to other situations. Next, we highlight issues in imaging fixed dipoles (i.e., fluorophores that emit radiation with an anisotropic angular distribution, producing noncircular images). Finally, we consider localizing fluorophores that are close enough to produce overlapping images. These techniques enable imaging of faster processes but are somewhat less mature. We also discuss localization algorithms in 3D super-resolution. Throughout, we identify practical guidelines for users and areas of potential improvement.

The principles of localizing single isotropic point sources

Estimation, precision and accuracy. Estimating a fluorophore position from an image is, in some sense, an exercise in geometry: without noise, an image of an isotropic light emitter would be a disk (possibly surrounded by diffraction rings) centered on the position of the fluorophore. Pixelation is only a minor complication: shifting a fluorophore by a fraction of a pixel (**Fig. 1**) causes detectable image asymmetry, with the degree of asymmetry depending on the

BOX 1 THE PSF AND THE GAUSSIAN APPROXIMATION

The point spread function (PSF) of an imaging system describes the shape of the blur formed when a point source is imaged; it is proportional to the average number of photons at a given position relative to the source. High-accuracy PSF calculations require accounting for numerous factors, particularly the collection angle (numerical aperture) of the lens, interfaces between the sample and lens (such as coverslips and immersion oil) and the dipole moment of the light source.

For an isotropic point source (emitting light equally in all directions), the most common models used in high-precision work are the Richards-Wolf model[99], which accounts for the vector nature of light waves, and the Gibson-Lanni model[100], which also accounts for coverslips and other interfaces between the sample and the lens. (An example image generated with the Richards-Wolf model is shown in **Fig. 2**.) As long as the lens is well-corrected for aberrations, the general pattern will be a central bright spot whose width corresponds roughly to the wavelength of light. However, although software is available for computing the Richards-Wolf and Gibson-Lanni PSFs (for example, an ImageJ plug-in[51]), the formulas are complicated. Many investigators therefore approximate the PSF with an Airy function, which is somewhat simpler for mathematical calculations[31,35]. The Airy PSF is most valid when the lens has a low numerical aperture, but it has sufficient qualitatively validity to be a useful approximation in many investigations.

Even the Airy PSF is still tedious for many practical calculations. Consequently, people often approximate the PSF of an isotropic source with a Gaussian function

$$I(x, y) = I_0 \exp(-a \times k^2((x - x_0)^2 + (y - y_0)^2)) + b \qquad (3)$$

where k is 2π divided by the wavelength of light in the specimen, a is a numerical factor that specifies the width of the PSF (often close to 0.25, though the precise value depends on the model being fit to[37]) and the other variables are defined as in equation (1). Both formulas are good approximations to more realistic PSFs and may thus be useful approximations in many situations (though the user is advised to make a careful choice of PSF width to ensure accuracy). Indeed, almost any realistic PSF will look roughly similar to a Gaussian bell curve near the peak of a focused image (though it can differ substantially out of focus, as discussed in the companion Review[28] and in ref. 54), which is the main reason why the Gaussian approximation gives useful and reasonably accurate results in many investigations involving focused images of fluorophores (for example, most of the ones cited in this Review). If, however, the image is defocused (a common situation in 3D), then the PSF shape can be considerably more complicated[51].

The Gaussian approximation remains useful in the presence of astigmatism if the exponent of equation (3) is modified to give an elliptical profile (i.e., different coefficients for the x and y terms)[40]. If, however, the pattern is asymmetric owing to dipole effects, random scattering or coma aberration, the Gaussian approximation is not necessarily valid. Additionally, in the tails, the approximation can break down (**Fig. 2**), as a Gaussian decays more rapidly than many PSFs. This poses issues in minimizing discrepancies between the model and the data in the edges of the image, so one must make a judicious choice of region-of-interest (ROI) size. Using a small ROI discards useful information, but a large ROI includes tails, where model mismatch effects may be substantial.

Another implementation issue is pixelation: the signal on a pixel is the sum of the photons striking different parts of it, which is proportional to the integral of the PSF over the pixel area. If the size of the pixel is substantially smaller than the PSF width (a fraction of a wavelength in the object plane, for a typical diffraction-limited PSF), the signal can usually be approximated with the PSF value at the center of the pixel. However, if the pixel is larger, or if high precision is required, the integral of the PSF over the pixel area must be used (for example, an error function for the Gaussian PSF).

distance moved. This asymmetry is the fundamental source of subpixel spatial information. Any automated procedure (which we shall commonly call an 'estimator') for determining fluorophore position from an image is, on some level, working from that subpixel information.

Real images contain several factors that complicate position estimation, the most important being shot noise (i.e., the signal is a random variable with a Poisson distribution). Photons are not distributed smoothly over the detector in exact accordance with the point spread function (PSF) of the system (**Box 1**; several example PSF shapes are illustrated in **Fig. 2**). Instead, they arrive at random positions; the PSF predicts only the mean number of photons on each pixel. As photon arrival positions are random, the image will deviate from the ideal PSF shape in a manner equivalent to adding noise to an 'ideal' image of the fluorophore. Consequently, each image will have a slightly different center, and any procedure for estimating the fluorophore's position will give different results for each image. The amount by which the position estimate varies among images is measured by the standard deviation of the estimates (as discussed in the Review by Deschout *et al.*[28]). The variation of the estimates is commonly referred to as precision, in keeping with common terminology for describing experimental data. A larger standard deviation implies less precision, as the estimates vary more widely.

Another confounding factor is background from out-of-focus fluorophores or scattered light. Besides adding noise to the image via fluctuations in the background, background shifts the 'average' photon position either to the center of the image (for uniform background, i.e., the same average number of background photons on each pixel) or toward the region of brightest background (for nonuniform background). In either case, the position estimates contain systematic errors, resulting in bias. The correction of uniform background is usually straightforward and is a key aspect of many algorithms discussed below. Correcting for nonuniform background is a harder problem than correcting for shot noise because nonuniform background will be (by definition) irregular and will depend on the particular distribution of out-of-focus sources and/or scatterers in the sample being studied. The most promising approaches thus far for nonuniform background still require that background be similar between subsequent frames and therefore would not correct for an out-of-focus fluorophore that blinks[29]. The PSF and noise distribution, on the other hand, tend to be more consistent across different samples. For lack of well-motivated models of nonuniform fluctuating background, this Review will mostly consider uniform background, which is easier to study and correct.

An issue related to background—one crucial to accuracy—is determining whether a camera frame has only one focused image of a fluorophore, or multiple focused (or nearly focused) images overlapping[21,30]. Position estimation from overlapping images is possible (see "Localization of fluorophores from multiple-fluorophore images"), but it requires an algorithm that can identify such images. Many researchers use analysis algorithms that work with only single-fluorophore images, in which case one must either remove the multiple-fluorophore images from the analysis or work under conditions in which only a very small fraction of the fluorophores is activated (to minimize the number of overlap images). Analyzing multiple-fluorophore images with an algorithm that assumes a single fluorophore will produce inaccurate or even meaningless results.

An additional issue is misspecification of the PSF. Any PSF model will be approximate at some level; even with a rigorous treatment

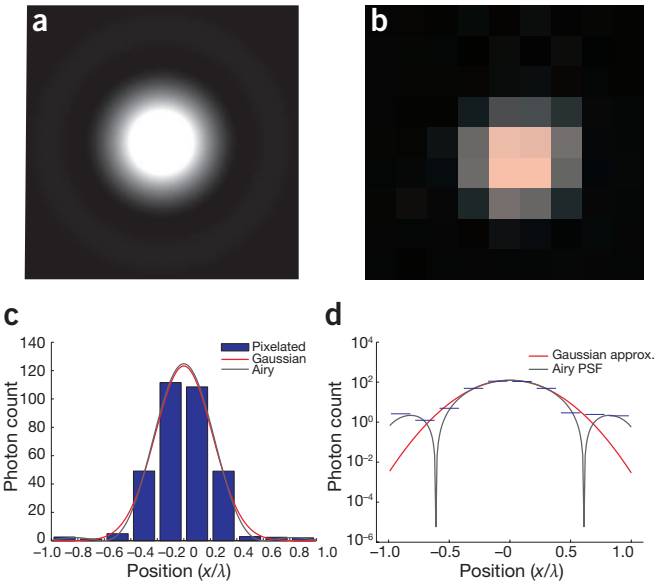

Figure 2 | Point spread functions. (**a**) Focused image of a fluorophore calculated from the Richards-Wolf model. The image is two wavelengths across. (**b**) The image in **a**, pixelated with one-fifth–wavelength pixels, and with shot noise (assuming 1,000 total photons in the image). (**c**) Cross-section of the pixelated image and approximate Airy and the Gaussian models for comparison. The widths of the model PSFs are based on standard theoretical models[37]. The heights are proportional to the photon count in the image. (**d**) Same as **c** but shown on a logarithmic scale to emphasize the discrepancy in the tails.

of lens aberrations and reflections at the lens surfaces, the inhomogeneity of the sample will introduce some amount of aberration. Nonetheless, some PSF models are more accurate than others. Misspecifying the shape of the PSF is known to affect localization precision in 2D (ref. 31) but generally does not introduce bias if the PSF is approximately symmetric, as is often true in 2D (ref. 32). However, many 3D imaging techniques require estimation of the width, shape or orientation of the PSF (in addition to its center in the 2D image), as do analysis techniques for dipoles. A misspecified PSF model might cause systematic biases in shape estimation and, thus, inaccuracy in the fluorophore's axial coordinate.

Image formation and noise models. The most mathematically rigorous approach to estimating a fluorophore's position is to treat it as a problem in statistics: one writes down a model for how the signal $I(x, y)$ on a pixel at (x, y) depends on the coordinates of the fluorophore, its photon emission rate, the background in the experiment and, perhaps, other parameters (such as aberrations or the depth of the source relative to the focal plane[33]). The parameters are varied to find the values that give the best 'fit' to the data, and fit is typically measured by either the maximum-likelihood criterion[31,34–36] or the least-squares (LS) criterion[15,16,18,37]. Once the best fit is found, coordinates are inferred from the fit parameters.

Fitting models are typically assumed to be of the form

$$I(x, y) = I_0 h(x - x_0, y - y_0) + b \quad (1)$$

where h is the PSF of the imaging system, I_0 is the peak intensity and is proportional to the photon emission rate and the single-frame

acquisition time, (x_0, y_0) are the fluorophore's coordinates, and b is the average background per pixel. Including background is important; failure to correct for background can bias estimates toward the center of the image, thereby decreasing accuracy. Some models assume a PSF whose width, shape and orientation depend on additional parameters: for example, the axial shift of the fluorophore relative to the focal plane[33,38–42]. Most fitting algorithms estimate all four (or more) parameters (x_0, y_0, I_0, b, etc.), but some estimate just the coordinates[16].

The difference between the measured signal and the model in equation (1) is due to a combination of model mismatch[31] (usually an inaccurate PSF) and noise. Model mismatch can be minimized by an intelligent choice of PSF, but noise is inevitable. The most common problem is shot noise, which is a result of the random nature of photon emission. Even if the detector is 100% efficient (i.e., every photon is successfully registered as an electronic signal), the number of photons emitted in a given time interval is typically a random variable following a Poisson distribution[43]. Consequently, the signal from the fluorophore and the out-of-focus background both follow Poisson distributions, as does their sum[34,36]. However, if one uses an electron-multiplying charge-coupled device (EMCCD), the electron multiplication process introduces additional noise[44–46]. To our knowledge, few existing algorithms explicitly account for multiplication noise, and one of the more notable algorithms to do so is optimized for low light levels[46]. Most work with EMCCD cameras has used localization algorithms that approximate the noise with Poisson statistics, often achieving excellent precision[47], but users should understand that such models are often only approximate.

The camera can also add 'read noise' (typically with a Gaussian distribution) in the process of converting an optical signal to an electrical signal. When the photon count from the in-focus fluorophore and background is low, Gaussian read noise may be non-negligible. The task of the user is to make an informed choice of noise model and PSF (whether an approximate formula or a measured profile[42,48]), use that information with a good localization algorithm, and interpret the results with appropriate caveats. Maximum-likelihood algorithms that take into account the noise characteristics of individual pixels in scientific complementary metal-oxide semiconductor (sCMOS) cameras are available[49] but require the user to experimentally characterize the noise in each pixel. When a key variable is not well known, the task is to either make a suitable approximation or pick an algorithm that can handle limited information.

Guidelines for evaluating a localization tool. Evaluating a localization algorithm and its PSF model usually requires testing with simulated data: creating an ensemble of images of the same fluorophore (each image distorted by noise), passing those images to the software and determining the mean and standard deviation of the position estimates. It may seem intuitive that the best test is with experimental data. However, unless one has a calibration sample in which the position and orientation of the fluorophore are known with high precision and accuracy, it is hard to distinguish between miscalibration of the sample and errors in the localization algorithm. Images can be simulated using a realistic PSF, realistic background and realistic noise, enabling tests in which the 'right answer' is known. Calibration standards have been developed on the basis of DNA rulers[4] and DNA origami[50], but simulated images remain important complements to experimental images when benchmarking localization algorithms. Also, the simulation's noise model should be matched to the camera, taking into account its quantum efficiency, additive read noise and (when appropriate) multiplication noise. The localization precision should be benchmarked against the theoretical limit to precision (the Cramér-Rao lower bound). A comprehensive understanding of the algorithm's performance requires testing on images with a range of photon counts from the fluorophore and also a range of background levels.

It is also important to remember that benchmarking the localization precision and accuracy of an algorithm is not the same as determining the resolution of an image produced by that algorithm. The precision and accuracy of a position estimate only tell the reliability of that particular fluorophore's estimated position. Image resolution also depends on whether the structure was labeled densely enough to resolve fine details (much as a ruler with 1-millimeter marks cannot be used to measure with 0.1-millimeter precision) and whether artifacts have been introduced during the analysis and reconstruction. Refining these ideas into information-theoretic measures is an important area of ongoing investigation[12–14]. A promising approach, taken from electron microscopy, is to examine different possible reconstructions using just half of the position estimates and seeing how similar these reconstructions are. The degree of similarity is determined by correlating the Fourier transforms of the images, producing a natural quantitative measure of resolution. The key lesson is that image resolution and localization precision are not synonymous and should not be confused when evaluating algorithm performance.

Localizing isotropic point sources via fitting

The most rigorous approach to localizing fluorophores from diffraction-limited image data is to fit a PSF model (of the form in equation (1)) to the data and vary the parameters to minimize mismatch between the data and model. It is usually straightforward to use an optimization routine to find such parameter values and obtain high-precision and (when background is included in the model) unbiased[31] results. The task for the user is to decide how to quantify mismatch. We describe here the two most common measures (LS and maximum likelihood estimation, MLE) and compare their strengths and weaknesses.

The least-squares criterion. LS fits require no detailed knowledge of the camera's noise, and they are common in other areas of data analysis. The error is the difference between the predicted signal and the actual signal, and one sums the weighted squares of the errors on all pixels.

$$S = \sum_{\text{pixels}} \frac{(\text{data}-\text{model prediction})^2}{\text{expected variance of data}} \qquad (2)$$

The parameters are varied to find values that minimize this weighted sum of squared errors (S), typically via the widely used and efficient Levenberg-Marquadt algorithm[31,51,52]. The weight in the denominator is the expected variance of the signal on that pixel. For Poisson noise, this variance is equal to the model's prediction for the signal on the pixel. Intuitively, weighting can be understood as comparing actual errors to expected errors. If the signal on a pixel is expected to have a large variance, then a large error does not necessarily reflect a poor fit; but if the expected variance is small, then even a small error may be a substantial problem. LS fits with weighting to the expected variance are mathematically equivalent to MLE (discussed below)

when the noise can be approximated as Gaussian. For high-background fluorescence (for example, ten or more photons per pixel), this is often a good approximation[53].

There are two important points to note when weighting data in LS. First, weighting gives extra importance to the tails of the PSF, where the expected signal (and variance) is low. If one has an accurate model for the tails of the PSF, weighting the tails is reasonable. If, however, the PSF is only approximate in the tails, errors that are small in absolute terms can be large in relative terms, owing to weighting by a small PSF. Consequently, some LS implementations do not use weighting[16,31,54,55]. When weighting is used most effectively, it is often with a PSF constructed from experimental data[42] (rather than an approximate formula). Model mismatch is not an issue with measured PSFs, provided that sample nonuniformity does not substantially alter the PSF and that it was measured from a calibration sample that could be shifted with subpixel precision. Misspecifying the PSF tail is also less of an issue when there is substantial background fluorescence such that the average background is comparable to or larger than the PSF tail[53,56].

Second, weighting should be done with respect to the expected variance (i.e., the model prediction), not a variance computed from measurements. Weighting to measured data means weighting to a combination of signal and noise; this runs counter to the goal of comparing a discrepancy between model and data with an expected baseline. Moreover, downward fluctuations in the data (due to noise) can result in small denominators, which may give greater weight to noisy pixels[53].

Maximum-likelihood estimation. This approach is prescribed by theorems[34,36] stating three key results. First, for any parameter that one might wish to estimate in statistics (for example, estimating the position of a molecule by analyzing a noisy image), there is a theoretical limit to the variance of unbiased estimates, or a maximum achievable precision. Second, when this maximum precision is attainable by an unbiased estimator (a condition that is not always assured), MLE will achieve it. Third, the variance of the maximum-likelihood estimate approaches the theoretical limit for a large data set.

MLE requires a model of signal (PSF) and noise (shot noise, for example, or shot noise plus Gaussian read noise). With this model, one computes the likelihood of obtaining the observed signal, assuming some estimated parameter values (x_0, y_0, I_0, b, etc.). Discrepancies larger than the typical noise level are unlikely; the likelihood of the data thus measures the fit between the data and the model. The parameters are varied to maximize the likelihood of the data. Because photons are usually independent of each other in imaging experiments, the likelihood of the data can be calculated by multiplying the likelihoods of the signals detected on each pixel. It is generally more convenient to work with the logarithm of the total likelihood, and the log likelihood is a natural measure of 'goodness of fit' between the data and the model[47]. A number of tools[31,47,52,53,56] perform MLE fits on single-fluorophore images.

MLE and localization precision. Besides estimating position with (often) the highest possible precision, a key advantage of MLE is that that precision is known. One can calculate the inverse of the Fisher information (roughly, the second derivative of the log likelihood)[34–36] and average it over all possible data sets (images) consistent with the model, which gives a prediction for the variance of the estimates. The inverse of this quantity is the expected variance of the parameter (position) estimates, and it has the name 'Cramér-Rao lower bound' (CRLB) after Cramér and Rao, who independently derived a fundamental limit to statistical estimation in the 1940s[57,58]. It is not always possible to estimate a parameter with the precision given by the CRLB, but when that degree of precision is achievable, the estimator that achieves it is MLE, and such high-precision estimation has been demonstrated in practice with realistic images[31,47]. Software is available to compute the CRLB for position estimates for given values of the pixel size, camera read noise, PSF width and shape, background level and photon count from the fluorophore[31]. A general guideline is that the standard deviation of the position estimates will be proportional to the PSF width (often the wavelength) divided by the square root of the number of photons detected.

The utility of the CRLB is twofold. First, in designing an experiment, one can estimate the best-case outcome for a given situation. Second, one can assess the performance of analysis software by feeding a representative ensemble of images to the software, obtaining position estimates and comparing the variance of the estimates with the CRLB.

Guidelines for choosing between LS and MLE. The greatest advantage of LS (especially unweighted LS) is that LS requires less detailed information on camera noise. Weighted LS requires knowing the variance of the noise on a given pixel but not a detailed probability distribution for the noise. Unweighted LS does not even require the variance, making it simple to use when noise is poorly characterized. (This is equivalent to doing MLE with a constant-variance Gaussian noise model.) MLE, on the other hand, requires a good understanding of one's camera, including gain, excess noise due to electron multiplication (for an EMCCD) and readout noise[44–46,49]. Shot noise has gotten the most attention because it is less mathematically tedious to incorporate than a model that combines shot noise and Gaussian read noise[35], but for many cameras it is not necessarily the most accurate model. Also, the variance of Gaussian read noise often differs between pixels, and the optimal way to address it in MLE is to incorporate pixel-by-pixel measurements of the noise[49]. In principle, one could do MLE without detailed knowledge of camera noise parameters, by including them as fit parameters, but it is dangerous to compensate for lack of knowledge by substantially increasing the number of adjustable parameters. The key point is that proper use of MLE requires a carefully measured noise model.

Beyond the issue of robustness against noise, several recent studies have compared LS and MLE in the presence of a number of complicating factors[31,47,53]. The most exhaustive comparison between MLE and LS is that of Abraham *et al.*[31]. As others have[47,53], they found that for a wide variety of photon counts and background levels, MLE gets closer to the CRLB, with LS typically being about 25% worse at low photon counts. At higher photon counts (whether due to brighter fluorophores or greater out-of-focus background), LS performs almost as well as MLE. The reason is that when the photon count is large, the shot noise has an approximately Gaussian distribution[34,36], and the LS criterion is mathematically equivalent to the maximum-likelihood criterion if the noise is Gaussian. However, at lower photon counts, shot noise cannot be well approximated with a Gaussian distribution, and adding Gaussian read noise to shot noise results in a very different noise distribution. Consequently, LS estimates may exhibit greater variance than MLE. Also, both approaches

can face problems from misspecification of the PSF, with MLE being more robust against misspecification of the PSF width and LS being more robust against misspecification of the PSF shape (for example, Airy vs. Gaussian)[31].

In both LS and MLE, PSF misspecification is most problematic in the tails because of small denominator problems. In LS, this arises from weighting: the expected signal is small in the tails, leading to a small denominator. In MLE, the formula for the log likelihood has an expected photon count in the denominator (see Smith *et al.*[47] for a particularly good derivation of this formula). This issue is most noticeable when fitting a Gaussian PSF (which decays rapidly in the tails) to an image generated with a PSF that decays more slowly in the tails. Model mismatch is especially problematic if the image is off-center in the region of interest (ROI)[56] because much of the data comes from the PSF's tail. The Gaussian PSF decays very rapidly, so the probability of detecting many photons in the tails is incredibly small. Consequently, even small fluctuations in the tail cause large relative changes in the likelihood of the data. This reinforces the need for a judiciously chosen ROI size, balancing the need for large photon count against mismatch in faint tails. An additional, simple safeguard is to recenter the ROI whenever the fitting routine estimates a fluorophore position more than one pixel from the center.

Both approaches can be implemented efficiently. A particularly fast implementation of LS is the Gaussian mask estimator[16], which varies only two parameters: the coordinates x_0 and y_0. A fast MLE implementation is also available, one that separately estimates x_0 and y_0 by first summing rows of pixels in the ROI and fitting the resulting data sums to a model for how the signal depends on the vertical coordinate y (giving a maximum-likelihood estimate of y_0) and then summing columns to do the same for x_0 (ref. 56). If the ROI size (in pixels) is $L \times L$, then each process works on only L inputs instead of L^2 inputs. MLE has also been implemented on graphics processing units (GPUs) for further speed improvements via parallelization[47,59]. Nonetheless, MLE sometimes takes longer to converge, especially if one varies the PSF width as a fit parameter (to account for defocus). The Newton-Raphson method[60] helps with fast convergence, but it is only a partial fix. An MLE algorithm adapted from the Levenberg-Marquadt LS algorithm has also been developed[61]. An additional safeguard is to start with a good position estimate—the closer that the initial parameter values are to the maximum-likelihood values, the fewer iterations needed for MLE to converge. A reasonable method for generating initial parameter estimates is to begin with LS[53]. LS converges quickly and generates an estimate that is (often) close to MLE.

In summary, one is well advised to favor MLE when adequate information is available on PSF shape and camera performance. Highly distorted PSFs, with complicated shapes due to scattering or aberrations, may be candidates for unweighted LS, as would imaging with a camera that exhibits large but poorly characterized fixed-pattern noise (i.e., performance varying pixel by pixel). These points are summarized in **Table 1**.

Table 1 | Comparison of the MLE and LS criteria for localization of single isotropic point sources

Maximum-likelihood estimation	Least-squares criterion
• Can, in principle, achieve theoretical limit of precision	• Often has lower precision but close to MLE precision for high photon counts and background
• Works best with a good model of camera noise	• Requires no information about noise; equivalent to MLE for Gaussian noise
• Requires a good PSF model for optimal performance but can use approximate PSF shape; PSF width can be a fit parameter	• Robust against misspecification of PSF shape but requires well-specified PSF width, or PSF width can be a fit parameter
• Takes more time to converge if PSF width is a fit parameter	
• Typically implemented with analytical PSF (i.e., a formula) but has been implemented with measured PSFs for 3D imaging[41,48]	
• Potential small-denominator problem when background is low and PSF tail is misspecified; this is solvable by proper centering and sizing of the ROI	
• Suitable for GPU implementation; fast algorithm available that fits x and y independently[56]	• Suitable for GPU implementation; fast algorithm available that fits only x and y and ignores other fit parameters[16]

Localizing isotropic point sources without fitting

Fitting is not the only way to localize fluorophores. Fitting requires evaluating a model on each pixel, comparing the model with data and then adjusting parameters and repeating the process. This can be cumbersome and requires (at least) a PSF model. Most alternatives to fitting attempt to be model independent, achieving both speed and broad applicability. Some of these algorithms are quite accurate and precise: no theorem or principle precludes an estimator from coming very close to the CRLB. Because of the prospect of high precision and fast, model-independent performance, alternative localization algorithms are of continuing interest.

Representative nonfitting algorithms. The simplest alternative to fitting is the centroid. If a photon is assumed to arrive at the center of the pixel, the average photon position, or centroid, can be computed quite easily, yielding a good estimate of fluorophore position. The principle is that slightly displacing the fluorophore within a pixel makes the corresponding side of the image slightly brighter, moving the average photon position (**Fig. 1**). Although this is not the universally best estimator, there are super-resolution analysis tools such as the popular QuickPALM software that use centroid estimates[62,63], and it has been shown to have advantages for localizing diffusing fluorophores, largely because of minimal assumptions about image shape[64]. At the very least, the centroid is a good first estimate for the beginning of an iterative fitting routine, but caution is required when it is used on its own.

Background biases the centroid toward the center of the image, giving a systematic error that does not average to 0. Consequently, centroid estimators require, at a minimum, correction for background. A good background-corrected centroid estimator, called virtual window center of mass (VWCM)[63], gives results competitive with LS fitting. VWCM utilizes the fact that the centroid is unaffected by background if the fluorophore happens to be perfectly centered in the ROI. VWCM iteratively redraws

the ROI, trimming off fractions of pixels if needed, to center the fluorophore until the centroid estimate converges to a stable value. Although not as precise as fitting, it is model independent and serves as a good first-pass estimate for a fitting routine.

Another model-independent approach uses the insight that the Fourier transform of a single-fluorophore image is the phase-shifted Fourier transform of the PSF[65]. Position estimation becomes a problem of phase estimation, which can be done in a single iteration after a fast Fourier transform. Although this technique is not as precise as fitting, many imaging and spectroscopy techniques move from space- or time-domain concepts to frequency-domain concepts as they mature; further innovations along these lines may be useful.

A particularly common alternative to fitting, which has been implemented in an integrated image analysis package[66], is fluoroBancroft, based on the principle of triangulation[19]. Because the signal on a pixel is determined by the distance to the fluorophore, one could (in principle) triangulate position using three pixels; to combat noise, one combines signals from more pixels. A model-independent algorithm is radial symmetry[32,67], valid when the PSF is circularly symmetric, so that contours of constant photon count are, neglecting noise and pixelation, circular. If one were to draw lines perpendicular to the surfaces of these circles, they would intersect at the center, i.e., the fluorophore position. The radial symmetry algorithms draw lines perpendicular to contours of constant signal and look for the point of closest intersection. These algorithms achieve performance close to the CRLB (and superior to fluoroBancroft) in a single iteration, even for many noncircular images.

When to choose a nonfitting method. In general, fitting methods (especially MLE) are best when accurate noise models and PSFs are available. When computational simplicity is crucial (for example, field-programmable gate array implementation on-chip[68]), single-iteration approaches may be preferable. Single-iteration estimates can also be used as starting values for an iterative fitting routine. Algorithms that make fewer detailed assumptions about the PSF shape (for example, VWCM[63] and radial symmetry[32,67]) may also be preferable for deep-tissue imaging with highly distorted PSFs. However, users are cautioned that nonfitting approaches are not usually designed for particular noise sources and might not approach the CRLB of localization precision in the presence of multiple noise sources.

Areas for further development. The key insight of radial symmetry algorithms is that constant-signal contours contain positional information. An intriguing question is whether similar insights could work for oriented dipoles. The PSF of a dipole is a combination of spherical harmonics of low order, raising the possibility that the contours have universal patterns that carry spatial information. Another issue is camera noise. Detailed noise models are generally associated with MLE approaches, but it is an open question whether camera-specific noise issues can be compensated for in nonfitting algorithms.

Nonuniform background has not been well explored in most studies. Uniform background is a reasonable first approximation of experimental conditions, and it is unlikely that any algorithm could compensate for background varying substantially on distances comparable to the PSF width. (Indeed, the definition of out-of-focus fluorophores is that their images do not vary over short distances.) However, performance in the presence of slowly varying background merits further study. So far the most promising approach to nonuniform-background estimation is to consider background only in the immediate vicinity of the activated fluorophore and to compare frames with and without the activated fluorophore[29]. This approach can correct for background that varies rapidly in space but not in time; the problem of background from defocused blinking fluorophores remains open. The Fourier-domain approach[65] also seems especially promising for this task.

Localizing single fluorophores with oriented dipoles

A circular PSF is a valid assumption if the fluorophore's dipole moment (the axis along which charge oscillates, analogous to an antenna) undergoes rotational diffusion during image acquisition, i.e., if the signal is averaged over many dipole orientations. The best empirical support for the isotropic assumption is simply that it usually works: the ability to localize fluorophores with precision close to the CRLB in many of the studies cited above, using experimental data and isotropic PSF models, suggests that in numerous situations the fluorophores are rotating enough to give approximately isotropic PSFs. However, not all dipoles rotate, and, as discussed by Deschout *et al.*[28], a dipole with fixed orientation will preferentially emit light in directions perpendicular to the dipole moment.

From an image-processing standpoint, the issue is models, not algorithms. Given an accurate model for the image formed by a dipole with a given position and orientation[69–71], one can fit that model to the image. Tools are available to implement those models in MLE[53,72] as well as LS[73]. However, displacing a dipole along the optic axis (i.e., perpendicular to the focal plane) can produce lateral shifts of the image centroid[22]. Thus, a good model is not always enough: one could get position estimates that are very precise (i.e., clustered in a narrow range) but also very inaccurate because of the systematic lateral shift. For high-precision imaging of fixed dipoles, it is thus advisable to combine a good model with an experimental approach that provides 3D information as discussed in the next section. The double-helix PSF method is particularly suitable, as image asymmetry carries information on dipole orientation while image orientation carries 3D information[74,75]. Polarization information can also be used[75,76], though a full analysis of the achievable resolution must take into account noise limits on polarization measurements[77]. Fundamental limits to polarization-sensitive super-resolution merit further study.

A harder problem is that a dipole that moves enough to partially 'smear out' the PSF but not so much that the PSF is fully isotropic. Such images should be very dependent on the fluorophore's local environment, so it isn't clear what sort of model to use. Work on this question is needed, particularly for rotational subdiffusion.

3D super-resolution

Whereas most 2D super-resolution experiments follow similar principles, 3D super-resolution experiments work from a variety of distinct principles, and in each case the estimation of the third coordinate (z) involves distinct image processing tasks. The companion Review by Deschout *et al.*[28] discusses hardware issues for these different implementations; we just note here that each 3D super-resolution approach produces images with distinct characteristics that require appropriate fitting models. In some cases, the extensions of 2D approaches are straightforward. For instance, in astigmatism-based imaging[24,40], fitting an elliptical PSF model is sufficient (wherein the eccentricity is a proxy for the z coordinate)

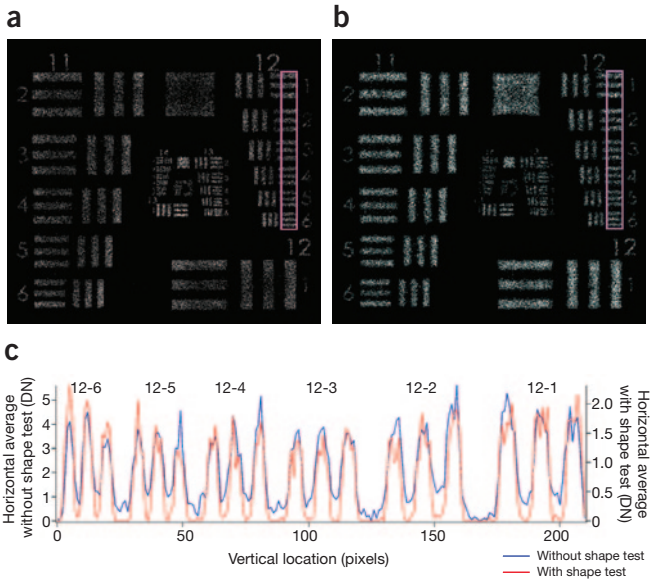

Figure 3 | Simulated localization microscopy images. (**a**) Simulated image of a US Air Force test pattern, assuming a dense labeling exponential distribution of photon counts. Localization was done with a fast MLE routine[56]. The image was reconstructed with multiple-fluorophore overlaps removed via a test of ellipticity. (**b**) Reconstruction without multiple-fluorophore overlaps removed. More points are visible in the reconstruction, but smaller features are not visible. (**c**) Horizontally averaged density of localizations (digital number per pixel, or DN) vs. vertical coordinate in boxed purple regions in **a** and **b**. Numbers above peaks refer to groups of bars on the edge of the test pattern.

and has been implemented on GPUs with precision close to the CRLB[47]. A tool for MLE with an elliptical PSF and data with many overlapping images is also available[78]. Biplane imaging[38] is even simpler from an image-processing standpoint, requiring just the fitting of two circular PSFs with variable width. Estimating z involves comparing the widths of two different images, which is feasible via LS fits to measured (rather than theoretical) PSFs, yielding results that compare favorably with astigmatism[42]. This approach can be generalized to more focal planes and more complicated PSFs[79,80].

Imaging with double-helix PSFs[41] requires fitting to a double-lobed PSF, and the orientation of the pair of lobes (vertical, horizontal or an intermediate diagonal) is directly proportional to the z coordinate (hence, 'helix' in the name). The PSF model is, naturally, more complicated than a 2D isotropic PSF, especially if one takes separate images in each polarization channel[75]. The CRLB for all three coordinates is better than in biplane or astigmatic imaging, and the localization precision is more robust against axial displacements[39,81]. One cautionary note is that when the PSF model has many parameters, there is the danger that an inaccurate estimate of one parameter might accidentally be compensated for by inaccuracy in another parameter because there are more ways to achieve a similar degree of match between data and model. When benchmarking an algorithm with several PSF parameters, one is thus cautioned to check the accuracy of all of the parameter estimates, not just the coordinates.

Although the approaches described here do not exhaust the list of 3D super-resolution microscopy techniques, they do cover most of the more popular methods, and many other 3D techniques[82,83] tend to have straightforward image processing requirements that scarcely differ from those of 2D techniques.

Single-fluorophore image analysis issues beyond position estimation

In addition to estimating fluorophore positions, super-resolution image analysis algorithms must identify an ROI and determine how many fluorophores are activated in the ROI.

ROI identification. Identifying bright spots is a common task. What is more particular to localization microscopy is determining whether the spot contains a single-fluorophore image[21,29], as single-fluorophore (or, increasingly, few-fluorophore) images are the key to super-resolution. The criteria for ROI identification determine the types of images passed for position estimation. At high signal-to-noise ratios (i.e., with photon count exceeding background fluctuations), essentially all single-fluorophore bright spots will be analyzed. However, low signal-to-noise ratios (relevant for thick specimens with substantial out-of-focus background) will give many marginal cases. Consequently, CRLB calculations based on averages over all possible images may not tell the whole story. One should compute the variance of the position estimates for the images that are actually passed to the position estimator.

Rejecting multiple-fluorophore images when using single-fluorophore models

Distinguishing single-molecule images from multiple-molecule overlaps has implications for the accuracy of image reconstruction (**Figs. 3** and **4**). In a densely labeled region with many fluorophores in close proximity, overlaps are highly probable. If one were to pass an overlap spot to a localization algorithm, the algorithm would typically estimate a position intermediate between the fluorophores that actually formed the image. A rejection step will reduce the density of such inaccurate localizations, as evidenced by a comparison of the reconstructions with and without a multiple-molecule rejection step (**Fig. 3**). The contrast between labeled and unlabeled regions is noticeably improved by rejecting overlaps with a shape test, as is evident from a visual comparison of the images as well as a quantitative comparison of localization densities. Moreover, without a shape test there may be substantial artifacts in the reconstruction (**Fig. 4**). Clearly, rejecting multiple-fluorophore images leads to a more accurate reconstruction, i.e. one with fewer artifacts.

Owing to the demonstrated importance of rejecting overlaps, most single-fluorophore localization algorithms have implicit or explicit criteria for identifying noncircular images. The criterion may be the image's ellipticity[4,56] (which is convenient if the test is done before position estimation) or the fit between data and model (which is common when the test is done after position estimation). In testing the fit between data and model, one is implicitly testing whether the image shows a single fluorophore.

Besides improving image quality, removing overlaps directly affects experimental design: increasing the activation probability per fluorophore increases the rate at which information can be acquired (temporal resolution) but also increases the risk of obtaining overlap images. If one can reject multiple-fluorophore overlaps, one can use higher activation probabilities to conduct faster experiments[21,29]. The relevant parameters are the acceptance probabilities for single-fluorophore and multiple-fluorophore images[54,84,85]. These probabilities are related to the recall and stochastic precision metrics[54]

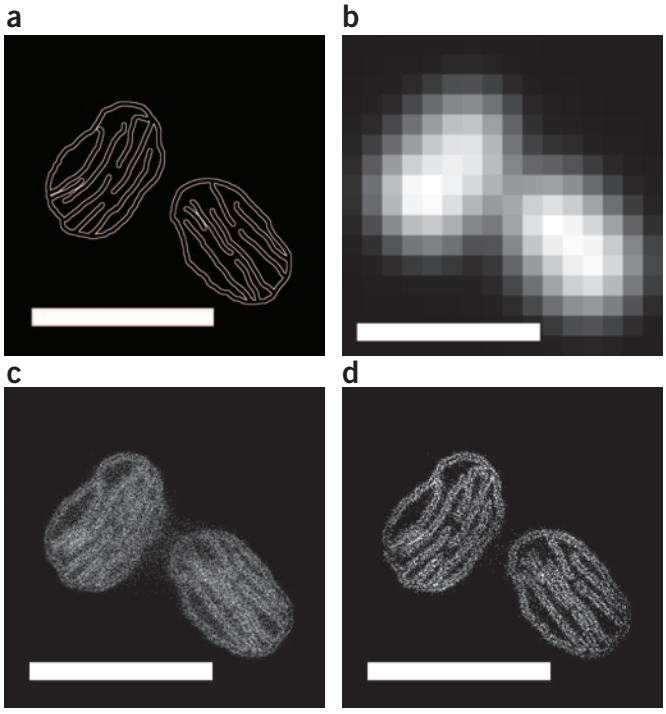

Figure 4 | Effects of accepting or rejecting multiple-molecule overlap images. (**a**) Structure used to generate the images. (**b**) Simulated diffraction-limited image if all fluorescent labels are simultaneously activated. (**c**) Reconstruction in which all bright spots are passed to an MLE algorithm. (**d**) Reconstruction in which only those bright spots passing an ellipticity test are included. Removing the elliptical spot makes folds more distinctive and removes artifacts. Scale bars, 500 nm (wavelength of light).

and can be computed by passing single-fluorophore and multiple-fluorophore images to analysis software.

Guidelines for users. Useful rejection criteria for overlap images depend in part on the shape and feature size of the fluorescently labeled structure. Consider, for instance, a linear cytoskeletal filament, a commonly imaged feature in localization microscopy[24] (**Fig. 5**). In order to be resolved, the filament must be densely labeled along its axis, with fluorophore separations shorter than the typical filament spacing (the Nyquist criterion[11]). Overlapping images from fluorophores located on different tubules will usually be highly elliptical and, hence, easy to reject. Conversely, overlaps of fluorophores located on the same tubule will often be from fluorophores separated by short distances and thus may be less elliptical and harder to reject. However, erroneously accepting these images will give positions that are still located on the same tubule; therefore, even these erroneous estimates will still be consistent with the structure. The key point is that rejection algorithms need to be benchmarked for their performance on multiple-fluorophore images with different fluorophore separations and distributions, and the relevant cases will depend on the expected structure of the specimen[23,29].

Fluorophore characteristics must also be taken into account. Photon count is not a perfect proxy for the number of activated fluorophores because the time in the activated state is a random variable[4]. However, a sufficiently large photon count is still a plausible indicator of a multiple-fluorophore image, enabling rejection of at least some images of closely spaced fluorophores and thus loosening performance tolerances for shape tests[29]. This requires adequate knowledge of how the fluorophore's quantum efficiency is affected by local environment. Moreover, assembling an image from fluorophore positions requires information on fluorophore blinking dynamics: if two frames have fluorophores at nearby positions, the details of the fluorophore's on-off process determine the probability that the frames show the same fluorophore[86].

Localization of fluorophores from multiple-fluorophore images

Some algorithms estimate positions from overlapping images of closely spaced fluorophores. Using multiple-fluorophore images enables one to activate more fluorophores simultaneously. Better image processing thus translates directly to a faster experiment and insights into processes on shorter time scales[49]. Algorithms for multiple-fluorophore images can be categorized by their outputs: either lists of positions or density profiles. The first category comprises straightforward extensions of algorithms discussed above. Algorithms in the second category typically give maps that are 'fuzzy' on length scales comparable to the localization precision.

Fitting-based methods. In these methods, one fits a multiple-fluorophore model to the data. The model is a sum of PSFs from fluorophores at different positions, and those positions are among the parameters varied to minimize mismatch between the data and model (**Fig. 5**). The most mathematically rigorous approach is, of course, MLE, and it has been implemented for multiple-fluorophore models by multiple groups[87], including GPU implementations[49,88,89]. Using MLE to obtain multiple fluorophore positions from an overlap image actually predates the synthesis of controllable blinking and localization that enabled super-resolution localization microscopy[90]. One particularly important implementation is highly optimized for the noise characteristics of sCMOS cameras[49]. An LS implementation (DAOSTORM) for multiple-fluorophore models is available[91] and is based on an algorithm from astronomy[92].

Many of the above comparisons of MLE and LS still apply: MLE can, with a reasonably accurate model, approach the theoretical minimum variance, but it tends to be computationally complex and requires a noise model. The most important performance benchmarks for multiple-fluorophore fitting algorithms are the localization precision, false positive rate (the number of position estimates that do not correspond to a fluorophore position) and localization

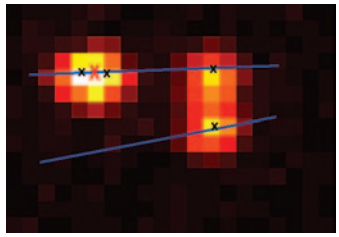

Figure 5 | Schematic comparison of single-fluorophore and multiple-fluorophore analysis. If we fit to a single-molecule model, only the top left pair of fluorophores (top left bright region) form an overlap image that is circular enough to be accepted by the algorithm. The position estimate (red X) is intermediate between the two fluorophores and is still on the fiber that they both label. The right pair form an ellipse that is not accepted by a single-fluorophore algorithm. However, if we fit to a multiple-fluorophore model, four position estimates (black Xs) can be obtained from this one frame.

fraction (the fraction of activated fluorophores correctly identified). MLE is known to achieve localization precision close to the CRLB[88,90]. In at least one implementation (PALMER[89]), MLE has a lower false positive rate and higher localization fraction than a common LS implementation (DAOSTORM[91]).

The multiple-fluorophore case also brings additional issues. An algorithm might favor a model with more fluorophores than are actually present, as adding more parameters to a model can often improve fit without conveying information on real phenomena. If the PSF width is a fit parameter (to account for defocus), then adding a fluorophore with a wide PSF and small photon emission rate might improve the fit by modeling chance nonuniformity in the background. It is thus necessary to pick parameters on the basis of more than just fit. Most algorithms stop adding fluorophores to the model when the residual or log likelihood reaches a user-determined threshold. If this criterion is somewhat arbitrary and ignores actual fluorophores, it at least ignores those making only small contributions to the signal.

Image estimation. Alternately, one can estimate a local density (or concentration) of fluorophores. In principle, this density should be 0 everywhere except at the position of a fluorophore. Density is estimated on a grid that is finer than the pixelated raw image (for example, dividing each pixel into 8 × 8 subpixels; refs. 93,94). In super-resolution, this task has been accomplished by borrowing tools from different fields, in particular Richardson-Lucy (RL) deconvolution[93,95,96] and compressed sensing (CS)[94]. RL deconvolution is an MLE approach to the problem of finding a fluorophore density map: for a given density map, PSF and noise model, one can

Table 2 | Summary of some openly available localization algorithms discussed in this article

	Fitting approach and PSF	Common implementations	Noise model	Notes for use
Single-fluorophore fits	MLE with isotropic PSF	Ober lab[a], Lidke lab (GPU implementation[47]), rapidSTORM[52], M2LE[56]	All assume shot noise; Ober's software also allows Gaussian camera read noise	Good for fluorophores with freely rotating dipole moments. Usually use a Gaussian PSF. Defocus can be accounted for via variable PSF width. M2LE includes an ellipticity test for rejection of multiple-fluorophore images
	MLE with elliptical PSF	Lidke lab, rapidSTORM	Shot noise	Most useful for astigmatism-based 3D imaging if the model assumes an ellipse oriented along one of the detector axes. Useful for rejection of two-molecule overlaps when the ellipse is arbitrarily oriented
	MLE with isotropic PSF, EMCCD excess noise and read noise	UAIM by Ober lab[46]	Combination of Poisson noise, electron-multiplication noise of EMCCD and Gaussian read noise	Optimized for use with very high magnification, but the noise model is applicable to almost any single-molecule experiment with an EMCCD
	LS with experimental PSF	Bewersdorf lab[41]	No detailed assumptions, but performance approaches theoretical limit if noise is a Gaussian; background correction is possible	Developed with particular attention to defocused fluorophores for 3D biplane imaging
	Fast LS with circular Gaussian PSF	Gaussian mask[16]		Practical when the PSF is not known in detail or when computational time is crucial
	Center of mass	Virtual window center of mass (VWCM)[63]	No detailed assumptions	Appropriate for diffusing fluorophores[64]. Good first-pass estimate to seed an iterative fitting routine. Designed for background correction
	fluoroBancroft[19]	LivePALM[66]	No detailed assumptions	Assumes a Gaussian PSF and requires single iteration
	Radial symmetry	Parthasarathy lab[32], Ma lab[67,68]	No detailed assumptions	PSF is only assumed to be radially symmetric. Performance is good even for nonradial PSFs[32]
Multiple-fluorophore fits	MLE with multiple-fluorophore model	Lidke lab[88], PALMER[87,89], Bewersdorf lab[49], 3D DAOSTORM[78]	Poisson noise in all implementations, sCMOS camera read noise in some implementations[49]	Assumes a Gaussian PSF and has fast GPU implementation
	LS fits to multiple-fluorophore models	DAOSTORM[91]	No explicit assumptions, but LS fits approach theoretical limit when noise is Gaussian	Good for imaging with a high density of activated fluorophores producing overlapping spots. PSF model allows for elliptical PSF found in 3D astigmatic imaging
Estimating a fluorophore density	Bayesian estimation	3B (ref. 97)	Poisson noise plus Gaussian read noise	Good for imaging a high density of activated fluorophores producing overlapping spots. Computationally intensive but can be sped up with cloud computing[98]
	Compressed sensing	Zhu lab and Huang lab[94]	No detailed assumptions in current implementations, but incorporating a detailed noise model is a direction for future work	Designed for imaging with a high density of activated fluorophores producing overlapping spots. Most useful if the PSF is the same for all fluorophores, but the PSF does not necessarily have to be circular
	Deconvolution	deconSTORM[93]	Poisson noise	Good for high-density images with multiple fluorophores producing overlapping spots. [a]http://www.wardoberlab.com/software/estimationtool/

This summary is non-exhaustive, and new tools are continually being developed. Readers are encouraged to look for tools that have been benchmarked by other users and successfully deployed in applications relevant to the reader's intended experiment.

compute the probability of obtaining that data set. The density map that maximizes the likelihood of the data set is chosen by RL deconvolution. Although it requires a noise model, it has the advantage of being able to incorporate information on the fluorophore blinking dynamics.

CS, on the other hand, requires no information on noise or blinking dynamics. It uses the fact that if one subdivides pixels, the number of unknown variables (i.e., the number of points in the density map) is greater than the number of input variables (i.e., the number of photon counts registered on pixels). The problem is thus underdetermined, having more variables than equations, and multiple density maps would be consistent with the same data. CS chooses the 'sparsest' map, i.e., the one for which the density is 0 on most grid points, reflecting prior knowledge that the specimen is labeled with discrete fluorophores. In direct comparisons on the same data, CS compares favorably with DAOSTORM. The density of fluorophores identified by CS is generally higher than in multiple-fluorophore MLE. Both CS and deconvolution are likely to continue to improve, owing to extensive experience with these techniques in other contexts.

Another interesting approach uses Bayesian statistics to compute the conditional probability of a density map[97]. This approach considers the entire sequence of images instead of separately analyzing each frame, and it takes into account fluorophore blinking dynamics. because multiple density maps may be similarly likely for a given data set, the final image is averaged over many density maps, sometimes revealing features not apparent in individual maps. This approach is computationally intensive, but it holds interesting possibilities for further development, and it can be sped up with cloud computing[98]. One might consider generalizing it to use prior spatial information to compute weighting probabilities. For instance, one would expect that images of the cytoskeleton will have Fourier transforms with peaks corresponding to the characteristic fiber spacing. Models with the expected peaks might justifiably receive greater weight. Similar considerations might apply to structures with known symmetry.

Guidelines for users. Tools using multiple-fluorophore models are not as mature as tools with single-fluorophore models, and the theoretical limits to localization with multiple-fluorophore models have not been as extensively studied. General guidelines are thus harder to give. As always, the user should use the best available PSF model. The parameter spaces being explored are larger, so real-time visualization may not always be feasible without skillful use of parallel processing. When possible, one should not activate too many fluorophores at once; even if the average number of activated fluorophores per ROI is within the capabilities of the localization algorithm, in dynamic samples the fluorescently labeled molecules or structures might aggregate (on time scales longer than a single frame), increasing the number of overlapping images in certain regions. Consequently, the activation probability should be chosen to account for the possibility that regions with low densities in early frames may have higher densities in later images.

Conclusions

The choice of a localization algorithm for super-resolution microscopy depends on what the user knows about the imaging system (PSF and noise), how the user controls fluorophore activation (i.e., single-fluorophore vs. multiple-fluorophore models) and whether the fluorophore's dipole moment is fixed. The gold standard for comparisons is MLE because it can often achieve the minimum variance in the position estimates, provided that the user possesses sufficient information. Alternative algorithms may be desirable for speed or for 'seeding' an MLE package with an initial estimate. **Table 2** provides a non-exhaustive summary of some of the algorithms discussed here. Issues meriting more exploration include camera noise, multiple-fluorophore fits in 3D, and distinguishing an image of a single oriented dipole from asymmetry created by the overlap of two images from different fluorophores. Given the reliance of super-resolution localization microscopy on good image processing, further advances in localization algorithms will likely further expand our experimental capabilities.

ACKNOWLEDGMENTS

A.S. was supported in part by a Teacher-Scholar Award from California State Polytechnic University. S.S. was supported by awards from the Microscopy Society of America and an award from the Kellogg Undergraduate Scholars Program of California State Polytechnic University.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Arch. Mikroskop. Anat.* **9**, 413 (1873).
2. Lidke, K., Rieger, B., Jovin, T. & Heintzmann, R. Superresolution by localization of quantum dots using blinking statistics. *Opt. Express* **13**, 7052–7062 (2005).
3. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
4. Rust, M.J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006).
5. Hess, S.T., Girirajan, T.P.K. & Mason, M.D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258–4272 (2006).
6. Gordon, M.P., Ha, T. & Selvin, P.R. Single-molecule high-resolution imaging with photobleaching. *Proc. Natl. Acad. Sci. USA* **101**, 6462–6465 (2004).
7. Qu, X., Wu, D., Mets, L. & Scherer, N.F. Nanometer-localized multiple single-molecule fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* **101**, 11298–11303 (2004).
8. Gustafsson, M.G.L. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. USA* **102**, 13081–13086 (2005).
9. Hell, S.W. Far-field optical nanoscopy. *Science* **316**, 1153–1158 (2007).
10. Westphal, V. & Hell, S.W. Nanoscale resolution in the focal plane of an optical microscope. *Phys. Rev. Lett.* **94**, 143903 (2005).
11. Shroff, H., Galbraith, C.G., Galbraith, J.A. & Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat. Methods* **5**, 417–423 (2008).
12. Nieuwenhuizen, R.P.J. *et al.* Measuring image resolution in optical nanoscopy. *Nat. Methods* **10**, 557–562 (2013).
This article borrows a technique from electron microscopy to formally quantify the resolution of a reconstructed image in localization microscopy.
13. Fitzgerald, J.E., Lu, J. & Schnitzer, M.J. Estimation theoretic measure of resolution for stochastic localization microscopy. *Phys. Rev. Lett.* **109**, 048102 (2012).
14. Mukamel, E.A. & Schnitzer, M.J. Unified resolution bounds for conventional and stochastic localization fluorescence microscopy. *Phys. Rev. Lett.* **109**, 168102 (2012).
15. Cheezum, M.K., Walker, W.F. & Guilford, W.H. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378–2388 (2001).
16. Thompson, R.E., Larson, D.R. & Webb, W.W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775–2783 (2002).
17. Högbom, J.A. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. Astrophys.* **15** (suppl.), 417–426 (1974).

18. Bobroff, N. Position measurement with a resolution and noise-limited instrument. *Rev. Sci. Instrum.* **57**, 1152–1157 (1986).
19. Andersson, S.B. Localization of a fluorescent source without numerical fitting. *Opt. Express* **16**, 18714–18724 (2008).
20. Bancroft, S. An algebraic solution of the GPS equations. *IEEE Trans. Aerosp. Electron. Syst.* **21**, 56–59 (1985).
21. Small, A.R. Theoretical limits on errors and acquisition rates in localizing switchable fluorophores. *Biophys. J.* **96**, L16–L18 (2009).
This article introduces a formalism for characterizing the performance of super-resolution image analysis algorithms and relating performance metrics to bounds on the quality of the reconstructed image. The article considers both single-fluorophore techniques and techniques that use multiple-fluorophore overlap images.
22. Engelhardt, J. *et al.* Molecular orientation affects localization accuracy in superresolution far-field fluorescence microscopy. *Nano Lett.* **11**, 209–213 (2011).
This article provides a particularly succinct and illuminating illustration of how dipole orientation complicates image analysis in single-molecule localization.
23. Fullerton, S., Bennett, K., Toda, E. & Takahashi, T. Camera simulation engine enables efficient system optimization for super-resolution imaging. in *Single Molecule Spectroscopy and Superresolution Imaging V* Vol. 8228 (eds. Enderlein, J., Gryczynski, Z.K., Erdmann, R., Koberling, F. & Gregor, I.) 822811 (SPIE, 2012).
24. Huang, B., Jones, S.A., Brandenburg, B. & Zhuang, X. Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. *Nat. Methods* **5**, 1047–1052 (2008).
25. Berro, A.J., Berglund, A.J., Carmichael, P.T., Kim, J.S. & Liddle, J.A. Super-resolution optical measurement of nanoscale photoacid distribution in lithographic materials. *ACS Nano* **6**, 9496–9502 (2012).
26. Lee, S.F., Thompson, M.A., Schwartz, M.A., Shapiro, L. & Moerner, W.E. Super-resolution imaging of the nucleoid-associated protein HU in *Caulobacter crescentus*. *Biophys. J.* **100**, L31–L33 (2011).
27. Wiltgen, S.M., Smith, I.F. & Parker, I. Superresolution localization of single functional IP3R channels utilizing Ca^{2+} flux as a readout. *Biophys. J.* **99**, 437–446 (2010).
28. Deschout, H. *et al.* Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods*, **11**, 253–266 (2014).
29. Hoogendoorn, E. *et al.* The fidelity of stochastic single-molecule super-resolution reconstructions critically depends upon robust background estimation. *Sci. Rep.* **4**, 3854 (2014).
29. Small, A.R. Theoretical limits on speed, errors, and resolution in microscopy with switchable fluorophores. in *Novel Techniques in Microscopy*, NMB4 (Optical Society of America, 2009).
31. Abraham, A.V., Ram, S., Chao, J., Ward, E.S. & Ober, R.J. Quantitative study of single molecule location estimation techniques. *Opt. Express* **17**, 23352–23373 (2009).
This article describes one of the more exhaustive comparisons of LS and MLE for single-molecule localization.
32. Parthasarathy, R. Rapid, accurate particle tracking by calculation of radial symmetry centers. *Nat. Methods* **9**, 724–726 (2012).
33. Deng, Y. & Shaevitz, J.W. Effect of aberration on height calibration in three-dimensional localization-based microscopy and particle tracking. *Appl. Opt.* **48**, 1886–1890 (2009).
34. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, 1993).
35. Ober, R.J., Ram, S. & Ward, E.S. Localization accuracy in single-molecule microscopy. *Biophys. J.* **86**, 1185–1200 (2004).
This paper introduces the key concepts of MLE and the CRLB for single-molecule localization. The derivations are particularly thorough and easy to follow.
36. Papoulis, A. & Pillai, S.U. *Probability, Random Variables, and Stochastic Processes* 4th edn. (McGraw-Hill, 2002).
37. Zhang, B., Zerubia, J. & Olivo-Marin, J.C. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.* **46**, 1819–1829 (2007).
38. Juette, M.F. *et al.* Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nat. Methods* **5**, 527–529 (2008).
39. Badieirostami, M., Lew, M.D., Thompson, M.A. & Moerner, W. Three-dimensional localization precision of the double-helix point spread function versus astigmatism and biplane. *Appl. Phys. Lett.* **97**, 161103 (2010).
This is one of the few applications of rigorous theory to quantify and compare performance of 3D super-resolution techniques. Most theoretical treatments have focused on 2D cases.
40. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).
41. Mlodzianoski, M.J., Juette, M.F., Beane, G.L. & Bewersdorf, J. Experimental characterization of 3D localization techniques for particle-tracking and super-resolution microscopy. *Opt. Express* **17**, 8264–8277 (2009).
42. Pavani, S.R.P. *et al.* Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. USA* **106**, 2995–2999 (2009).
43. Loudon, R. *The Quantum Theory of Light* 3rd edn. (Oxford Univ. Press, 2000).
44. Huang, Z.-L. *et al.* Localization-based super-resolution microscopy with an sCMOS camera. *Opt. Express* **19**, 19156–19168 (2011).
45. Quan, T., Zeng, S. & Huang, Z.L. Localization capability and limitation of electron-multiplying charge-coupled, scientific complementary metal-oxide semiconductor, and charge-coupled devices for superresolution imaging. *J. Biomed. Opt.* **15**, 066005 (2010).
46. Chao, J., Ram, S., Ward, E.S. & Ober, R.J. Ultrahigh accuracy imaging modality for super-localization microscopy. *Nat. Methods* **10**, 335–338 (2013).
The technique described in this paper is highly optimized for getting the best possible performance out of an EMCCD camera. Although the specific experimental approach may not be ideal for all situations, the insights into noise in EMCCD cameras are valuable for anybody working with them.
47. Smith, C.S., Joseph, N., Rieger, B. & Lidke, K.A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Methods* **7**, 373–375 (2010).
Besides describing a very useful GPU implementation of MLE, this article includes particularly useful derivations of key formulas related to MLE and the CRLB.
48. Quirin, S., Pavani, S.R.P. & Piestun, R. Optimal 3D single-molecule localization for superresolution microscopy with aberrations and engineered point spread functions. *Proc. Natl. Acad. Sci. USA* **109**, 675–679 (2012).
49. Huang, F. *et al.* Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods* **10**, 653–658 (2013).
The algorithm described here is optimized for the noise characteristics of sCMOS cameras and extends MLE to the case of overlap images with additive camera noise.
50. Schmied, J.J. *et al.* Fluorescence and super-resolution standards based on DNA origami. *Nat. Methods* **9**, 1133–1134 (2012).
51. Kirshner, H., Aguet, F., Sage, D. & Unser, M. 3-D PSF fitting for fluorescence microscopy: implementation and localization applications. *J. Microsc.* **249**, 13–25 (2013).
52. Wolter, S. *et al.* rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Methods* **9**, 1040–1041 (2012).
53. Mortensen, K.I., Churchman, L.S., Spudich, J.A. & Flyvbjerg, H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat. Methods* **7**, 377–381 (2010).
54. Wolter, S., Endesfelder, U., van de Linde, S., Heilemann, M. & Sauer, M. Measuring localization performance of super-resolution algorithms on very active samples. *Opt. Express* **19**, 7020–7033 (2011).
A key contribution of this article is its demonstration of how one can directly assess key performance metrics of a super-resolution image analysis algorithm and use those metrics to compare different algorithms.
55. Wolter, S. *et al.* Real-time computation of subdiffraction-resolution fluorescence images. *J. Microsc.* **237**, 12–22 (2010).
56. Starr, R., Stahlheber, S. & Small, A. Fast maximum likelihood algorithm for localization of fluorescent molecules. *Opt. Lett.* **37**, 413–415 (2012).
57. Cramér, H. *Mathematical Methods of Statistics* (Princeton Univ. Press, 1946).
58. Rao, C.R. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37**, 81–91 (1945).
59. Quan, T. *et al.* Ultra-fast, high-precision image analysis for localization-based super resolution microscopy. *Opt. Express* **18**, 11867–11876 (2010).
60. Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P *Numerical Recipes: The Art of Scientific Computing* 3rd edn. (Cambridge Univ. Press, 2007).
61. Laurence, T.A. & Chromy, B.A. Efficient maximum likelihood estimator fitting of histograms. *Nat. Methods* **7**, 338–339 (2010).
62. Henriques, R. *et al.* QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods* **7**, 339–340 (2010).
63. Berglund, A.J., McMahon, M.D., McClelland, J.J. & Liddle, J.A. Fast, bias-free algorithm for tracking single particles with variable size and shape. *Opt. Express* **16**, 14064–14075 (2008).
64. Deschout, H., Neyts, K. & Braeckmans, K. The influence of movement on the localization precision of sub-resolution particles in fluorescence microscopy.

J. Biophotonics **5**, 97–109 (2012).
65. Yu, B., Chen, D., Qu, J. & Niu, H. Fast Fourier domain localization algorithm of a single molecule with nanometer precision. *Opt. Lett.* **36**, 4317–4319 (2011).
66. Hedde, P.N., Fuchs, J., Oswald, F., Wiedenmann, J. & Nienhaus, G.U. Online image analysis software for photoactivation localization microscopy. *Nat. Methods* **6**, 689–690 (2009).
67. Ma, H., Long, F., Zeng, S. & Huang, Z.L. A fast and precise algorithm based on maximum radial symmetry for single molecule localization. *Opt. Lett.* **37**, 2481–2483 (2012).
68. Ma, H., Kawai, H., Toda, E., Zeng, S. & Huang, Z.-L. Localization-based super-resolution microscopy with an sCMOS camera part III: camera embedded data processing significantly reduces the challenges of massive data handling. *Opt. Lett.* **38**, 1769–1771 (2013).
69. Enderlein, J. Theoretical study of detection of a dipole emitter through an objective with high numerical aperture. *Opt. Lett.* **25**, 634–636 (2000).
70. Stallinga, S. & Rieger, B. Accuracy of the Gaussian point spread function model in 2D localization microscopy. *Opt. Express* **18**, 24461–24476 (2010).
71. Török, P., Higdon, P.D. & Wilson, T. Theory for confocal and conventional microscopes imaging small dielectric scatterers. *J. Mod. Opt.* **45**, 1681–1698 (1998).
72. Stallinga, S. & Rieger, B. Position and orientation estimation of fixed dipole emitters using an effective Hermite point spread function model. *Opt. Express* **20**, 5896–5921 (2012).
73. Aguet, F., Geissbühler, S., Märki, I., Lasser, T. & Unser, M. Super-resolution orientation estimation and localization of fluorescent dipoles using 3-D steerable filters. *Opt. Express* **17**, 6829–6848 (2009).
74. Backlund, M.P. *et al.* Simultaneous, accurate measurement of the 3D position and orientation of single molecules. *Proc. Natl. Acad. Sci. USA* **109**, 19087–19092 (2012).
75. Pavani, S.R.P., DeLuca, J.G. & Piestun, R. Polarization sensitive, three-dimensional, single-molecule imaging of cells with a double-helix system. *Opt. Express* **17**, 19644–19655 (2009).
76. Gould, T.J. *et al.* Nanoscale imaging of molecular positions and anisotropies. *Nat. Methods* **5**, 1027–1030 (2008).
77. Foreman, M.R. & Török, P. Information and resolution in electromagnetic optical systems. *Phys. Rev. A* **82**, 043835 (2010).
78. Babcock, H., Sigal, Y.M. & Zhuang, X. A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Opt. Nanoscopy* **1**, 6 (2012).
79. Ram, S., Prabhat, P., Chao, J., Ward, E.S. & Ober, R.J. High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells. *Biophys. J.* **95**, 6025–6043 (2008).
80. Ram, S., Prabhat, P., Ward, E.S. & Ober, R.J. Improved single particle localization accuracy with dual objective multifocal plane microscopy. *Opt. Express* **17**, 6881–6898 (2009).
81. Grover, G., Pavani, S.R.P. & Piestun, R. Performance limits on three-dimensional particle localization in photon-limited microscopy. *Opt. Lett.* **35**, 3306–3308 (2010).
82. Shtengel, G. *et al.* Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proc. Natl. Acad. Sci. USA* **106**, 3125–3130 (2009).
83. Tang, J., Akerboom, J., Vaziri, A., Looger, L.L. & Shank, C.V. Near-isotropic 3D optical nanoscopy with photon-limited chromophores. *Proc. Natl. Acad. Sci. USA* **107**, 10068–10073 (2010).
84. Shore, E. & Small, A. Optimal acquisition scheme for subwavelength localization microscopy of bleachable fluorophores. *Opt. Lett.* **36**, 289–291 (2011).
85. Small, A. Model of bleaching and acquisition for superresolution microscopy controlled by a single wavelength. *Biomed. Opt. Express* **2**, 2934–2949 (2011).
86. Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U. & Radenovic, A. Identification of clustering artifacts in photoactivated localization microscopy. *Nat. Methods* **8**, 527–528 (2011).
87. Quan, T. *et al.* High-density localization of active molecules using Structured Sparse Model and Bayesian Information Criterion. *Opt. Express* **19**, 16963–16974 (2011).
88. Huang, F., Schwartz, S.L., Byars, J.M. & Lidke, K.A. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. *Biomed. Opt. Express* **2**, 1377–1393 (2011).
89. Wang, Y., Quan, T., Zeng, S. & Huang, Z.-L. PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy. *Opt. Express* **20**, 16039–16049 (2012).
90. Ram, S., Ward, E.S. & Ober, R.J. Beyond Rayleigh's criterion: a resolution measure with application to single-molecule microscopy. *Proc. Natl. Acad. Sci. USA* **103**, 4457–4462 (2006).
91. Holden, S.J., Uphoff, S. & Kapanidis, A.N. DAOSTORM: an algorithm for high- density super-resolution microscopy. *Nat. Methods* **8**, 279–280 (2011).
92. Stetson, P.B. DAOPHOT: a computer program for crowded-field stellar photometry. *Publ. Astron. Soc. Pac.* **99**, 191–222 (1987).
93. Mukamel, E.A., Babcock, H. & Zhuang, X. Statistical deconvolution for superresolution fluorescence microscopy. *Biophys. J.* **102**, 2391–2400 (2012).
94. Zhu, L., Zhang, W., Elnatan, D. & Huang, B. Faster STORM using compressed sensing. *Nat. Methods* **9**, 721–723 (2012).
95. Lucy, L.B. An iterative technique for the rectification of observed distributions. *Astron. J.* **79**, 745 (1974).
96. Richardson, W.H. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* **62**, 55–59 (1972).
97. Cox, S. *et al.* Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nat. Methods* **9**, 195–200 (2012).
98. Hu, Y.S., Nan, X., Sengupta, P., Lippincott-Schwartz, J. & Cang, H. Accelerating 3B single-molecule super-resolution microscopy with cloud computing. *Nat. Methods* **10**, 96–97 (2013).
99. Richards, B. & Wolf, E. Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **253**, 358–379 (1959).
100. Frisken Gibson, S. & Lanni, F. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *J. Opt. Soc. Am. A* **9**, 154–166 (1992).

Corrigendum: Fluorophore localization algorithms for super-resolution microscopy

Alex Small & Shane Stahlheber
Nat. Methods **11**, 267–279 (2014); **published online 27 February 2014; corrected after print 23 July 2014**

In the version of this article initially published, 3D DAOSTORM was erroneously characterized in Table 2 as using least-squares fits. In fact, 3D DAOSTORM uses an implementation of maximum likelihood estimation (MLE). The error has been corrected in the HTML and PDF versions of the article.