Published in final edited form as:

Curr Opin Struct Biol. 2014 October ; 0: 112–121. doi:10.1016/j.sbi.2014.08.008.

Quantitative analysis of single-molecule superresolution images

Carla Coltharp^a, Xinxing Yang^a, and Jie Xiao^a

^aDepartment of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Abstract

This review highlights the quantitative capabilities of single-molecule localization-based superresolution imaging methods. In addition to revealing fine structural details, the molecule coordinate lists generated by these methods provide the critical ability to quantify the number, clustering, and colocalization of molecules with 10 – 50 nm resolution. Here we describe typical workflows and precautions for quantitative analysis of single-molecule superresolution images. These guidelines include potential pitfalls and essential control experiments, allowing critical assessment and interpretation of superresolution images.

Introduction

Light microscopy has profoundly increased our understanding of cell biology. The detail in this understanding, however, has long been limited by the diffraction of light [1]. In the last decade, superresolution fluorescence microscopy techniques have overcome this diffraction limit and substantially improved our ability to understand cell biology. By optically switching different subpopulations of molecules on or off at different times, these techniques temporally separate spatially overlapped molecules to achieve 10 – 50 nm spatial resolution [2]. The history, theory, instrumentation, and applications of superresolution microscopy have been excellently described elsewhere [3–5].

The present review will focus on quantitative data analysis of one major class of superresolution fluorescence imaging: single-molecule localization-based methods such as photoactivated localization microscopy (PALM) [6], stochastic optical reconstruction microscopy (STORM) [7], and their derivatives [8,9]. These single-molecule methods are unique because they provide lists of molecule coordinates in addition to intensity-based superresolution images. Proper analysis of the molecule coordinates enables quantitative characterization of the density, size, composition, and spatial distributions of cellular structures at near-molecular precision. Such quantitative information is difficult to obtain by other imaging techniques.

© 2014 Elsevier Ltd. All rights reserved.

Corresponding author: Jie Xiao (xiao@jhmi.edu; 1-410-614-0338).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Here, we first describe principles of single-molecule localization-based superresolution imaging, with a focus on spatial and temporal resolutions. We then provide a workflow of data analysis detailing how to quantify structural properties from the imaging data. We highlight caveats and pitfalls that affect data interpretation, and describe crucial control experiments and validations from notable recent experiments. As with any emerging technique, the best practices of superresolution image analysis continue to evolve, and we conclude by discussing some areas of active research in this development.

Principles of single-molecule localization-based superresolution imaging

Concept

The resolution of light microscopy is limited to ~ 250 nm because fluorescence from a single point source is blurred as light diffracts along the optical path of a microscope [1]. Single-molecule localization-based superresolution imaging methods circumvent this limit by ensuring that no more than a single molecule within a diffraction-limited area fluoresce simultaneously. This is achieved by using photoactivatable or photoswitchable fluorophores that can be switched on or off stochastically. Although fluorescence from a single molecule is still blurred by diffraction, the molecule can be localized with nanometer precision by estimating the centroid position of its fluorescence profile, termed the 'point spread function' (PSF). Figure 1A compares the diffraction-limited PSF of a molecule with its localization precision to illustrate the improvement of resolution (see Workflow Section 1). A superresolution image is generated by accumulating and superimposing hundreds to tens of thousands of molecule localizations (see Workflow Section 2). The localization list can then be analyzed to quantify different structural properties or spatial colocalization between species (see Workflow Section 3).

Spatial and temporal resolutions

The spatial resolution of a localization-based superresolution image is determined by both the precision of localizing each molecule, σ_{loc} , and the Nyquist resolution, σ_{Nyq} , which is related to sampling frequency within a structure (Figure 1B) [10]. Theoretically, σ_{loc} is determined by the signal-to-noise ratio and camera pixel size (Figure 1A, left equation) [11,12], and can be improved by increasing the number of photons collected from each molecule (Figure 1B, black dashed line). The true, experimental σ_{loc} is often lower than this theoretical value due to inefficient signal collection and thermal fluctuations, and can be measured in two ways. When using organic fluorophores that blink (switch on and off) many times, σ_{loc} can be calculated as the standard deviation of many fitted positions of the same molecule [13] (Figure 1C, left). For fluorophores such as fluorescent proteins that last only a few frames, σ_{loc} can be inferred from the pair-wise distribution of distances between sequential observations of all molecules [14,15] (Figure 1C, right).

The effect of sampling frequency on spatial resolution can be described by the Nyquist-Shannon sampling theorem, which states that to achieve a resolution of σ_{Nyq} , sampling with intervals smaller than $\sigma_{Nyq}/2$ is required [10,16]. Thus, higher sampling frequency improves spatial resolution. The effective spatial resolution of an image, σ_{img} , is limited by precisions in both localization and sampling, and can be described by $\sigma_{img} = \sqrt{\sigma_{loc}^2 + \sigma_{Nyq}^2}$ [17]. As

such, the ability to resolve well-sampled structures is limited by the localization precision (Figure 1D, i), whereas images generated with excellent localization precision can fail to resolve structural details due to undersampling (Figure 1D, ii) [10,18]. Improved resolving power requires optimization of both factors (Figure 1D, iii).

The temporal resolution of localization-based superresolution imaging is typically on the order of seconds to minutes. This relatively slow temporal resolution stems from the requirement that thousands of frames are needed to construct a superresolution image with reasonable sampling frequency. Due to this constraint, initial superresolution imaging was performed on fixed cells [19], and live-cell imaging was often performed on cellular structures with relatively slow dynamics [10]. Temporal resolution can be improved to around a second by increasing acquisition speed and by applying algorithms to distinguish molecules with overlapping PSFs in the same frame [20–25], but result in somewhat reduced localization precision (40 – 100 nm).

Workflow

1. Single molecule localization

Robust and efficient localization of single molecules is the first step of superresolution image construction and analysis. All localization methods involve estimating the centroid position of a single molecule's PSF, which is usually well-approximated by a 2D Gaussian function [11,26]. Under very high signal-to-noise ratio conditions, rotationally-fixed molecules may exhibit asymmetric PSFs that require alternate fitting functions [12,27]. Localization algorithms typically optimize either precision (iterative fitting methods) or computational speed (non-iterative methods). We refer the reader to excellent comparisons of localization methods [28,29], and methods to correct for systematic errors caused by uneven backgrounds [30,31]. A comprehensive list of localization software packages maintained by the Biomedical Imaging Group in Ecole Polytechnique Fédérale de Lausanne can be found online (<http://bigwww.epfl.ch/palm/software/>).

2. Image-based analyses

2.1 Rendering images—The output data of single-molecule localization algorithms contain molecule coordinates and their associated localization precisions, which are used to render a superresolution image. An intuitive method to represent this data is by superimposing unit Gaussian distributions centered on each molecule's coordinates. These Gaussian distributions are analogous to the PSFs that cause the familiar blur in conventional images (Figure 2Ai), but with standard deviations specified by each molecule's estimated localization precision (Figure 2Aii). Because the localization precision represents the uncertainty of each molecule's position, the intensities in this accumulated image represent relative probabilities that the detected molecules are located in each pixel but not the absolute number of molecules.

An alternative representation that allows comparison of molecule counts is a two-dimensional histogram of number of molecules per pixel (Figure 2Aiii). Molecules can only be confidently assigned to a particular pixel in this histogram if the pixel size is larger than the typical localization precision. Quantifying the absolute number of molecules in a

structure from this histogram requires correction for over- and under-counting artifacts (Section 3.1). However, these artifacts apply universally to the entire image, so relative molecule counts within different regions and features can still be measured directly from a sufficiently-sampled histogram without further correction [14].

2.2 Measuring structural dimensions—The dimension of a cellular structure is commonly measured from its cross-sectional projection [13,19,21]. With sufficient sampling frequency, this apparent cross-section is the result of the convolution of the true structural profile and the Gaussian distribution defined by localization precision (Figure 2B) [13]. Prior knowledge of structural shape can thus be used to deconvolve the true structural dimension from the apparent cross-section profile (see examples in Figure 2Bi and ii). In the absence of prior structural knowledge, the underlying structure can be approximated by a Gaussian-distribution (Figure 2Biii and iv), in which case the measured apparent full-width at half-maximum ($FWHM_{app}$) is related to the true dimension ($FWHM_{true}$) and the

localization precision ($FWHM_{loc} = 2.35\sigma_{loc}$) by $FWHM_{app} = \sqrt{FWHM_{loc}^2 + FWHM_{true}^2}$. Generally, dimensions of previously-uncharacterized structures are reported as fitted $FWHM_{app}$ because it allows convenient comparison of different structures observed with similar localization precision without imposing an assumption of structural shape [19,32]. As illustrated in Figure 2C, the deviation of $FWHM_{app}$ from the true dimension is most apparent for dimensions comparable or smaller than $FWHM_{loc}$, while measurements of larger structures are closer to their true values.

It is important to note that an additional systematic error is introduced when images are rendered by superimposing Gaussian distributions (Figure 2Aii, Section 2.1). This rendering method effectively convolves the localization precision twice and further increases apparent structural dimensions. However, this effect is only evident for structures having sizes comparable to or smaller than $FWHM_{loc}$ (Figure 2C, blue curve) [33]. Furthermore, because fitting the Gaussian-blurred image is more robust to factors such as sampling frequency and pixel size than fitting a molecule density histogram (Figure 2C, insets), dimensions are often still measured from the Gaussian-blurred image.

3. Coordinate-based analyses

Coordinate-based analysis is a valuable feature of single-molecule localization-based superresolution microscopy, which sets it apart from other superresolution approaches. We describe several methods to characterize spatial distributions patterns (Section 3.2) and colocalization (Section 3.3). Many of these methods are adapted from ecological analyses [34]. These methods require fixed or immobile molecules to ensure that molecule distributions do not change during data acquisition. Absolute molecule counting applications further require that the obtained coordinate lists are first corrected for over- and undercounting artifacts (Section 3.1).

3.1 Counting molecule numbers

Correcting for overcounting due to fluorophore blinking: One important prerequisite for absolute molecule counting is to correct for multiple observations of the same molecule due

to fluorophore blinking. Blinking results in overcounting and false clustering of molecules [35]. All current methods that correct for blinking require characterization of fluorophore properties using a monomeric, sparsely-distributed control sample ($\text{Control}_{\text{mon}}$) imaged under the same optical and chemical conditions as experimental samples [14,36].

A straightforward method to alleviate blinking artifacts is to identify spots localized close in space and time as those originating from a single molecule [14,35,37–39]. The distance threshold used to group spots in this manner should be based on the typical localization precision, and the time threshold should be based on the average ‘off’ time between successive blinks (characterized from $\text{Control}_{\text{mon}}$) [14,35,37,38**]. It is important to apply a sufficiently low activation level such that generous threshold values group most blinking events but do not incorrectly combine nearby molecules that happen to be activated in quick succession [14].

If the spot grouping method cannot be applied (e.g. when fast activation is required), true molecule clustering can be distinguished from blinking-related self-clustering using a pair-correlation function (PCF) [40**]. The PCF describes the pair-wise distance correlation between all localized spots (Section 3.2), which is the sum of the true spatial correlation function and the self-correlation function caused by blinking. The self-correlation PCF can be measured from the control sample ($\text{Control}_{\text{mon}}$) and subtracted from the experimental PCF to characterize true molecule clustering [40**].

Finally, the absolute number of detected molecules in a region can be estimated by dividing the number of localized spots by the average number of localizations per molecule (determined from $\text{Control}_{\text{mon}}$) without applying other blinking corrections [14,32]. This method only corrects mean molecule numbers and does not remove false clustering, so should not be used for analysis of spatial distribution patterns or clustering (Section 3.2).

Correcting for undercounting due to inefficient molecule detection: Inefficient fluorophore labeling and detection lead to undercounting of molecules in a structure. Because antibody binding often results in large variability in labeling efficiency, studies concerning absolute stoichiometry of complexes often use genetic labels such as fluorescent protein or affinity tag fusions [38**]. The labeling efficiency of genetic fusions can approach unity, but detection efficiency remains a major concern with any fluorescent label [41,42*]. To calibrate for detection efficiency, a tandem dimer of two fluorophores, either the same or of different colors, can be constructed [38**,39,43–45]. The construct should not self-associate and should be imaged under the same condition as the experimental samples. Assuming identical detection efficiency, p , for each fluorophore, the frequency of singularly-detected dimers (N_s) and dually-detected dimers (N_d) follows a binomial distribution [38**,42*]. The detection efficiency can be estimated from the ratio $N_d/N_s = p^2/2p(1 - p)$ or by fitting to a binomial distribution [38**]. Then, the true number of molecules in a structure can be determined by dividing the observed number of molecules in a structure by the detection efficiency.

3.2 Identifying spatial distribution patterns—A list of corrected molecule coordinates can be used to characterize the spatial arrangement or clustering of detected molecules. In

general, two classes of methods, correlation- or threshold-based, are used. Correlation-based methods use unbiased statistical analyses to describe overall clustering of molecules in an image. Threshold-based methods use experimenter-defined thresholds to segment clusters in an image and hence allow characterization of individual cluster properties.

Quantification of average cluster properties using correlation-based analyses: One common correlation-based method is the aforementioned pair-correlation function (PCF). The PCF calculates the probability of one molecule appearing at a certain distance r from another one [46] and can be computed using 2D FFT (fast Fourier transform) [40**]. For randomly distributed molecules, the PCF curve is constant (related to the density of molecules) at different r (Figure 3Ci). For molecules distributed in clusters, the probability of finding molecules within clusters (smaller r) is higher than that outside (larger r), and hence the PCF curve has higher values at short distances and decays at longer distances (Figure 3Ci, blue and purple lines). Theoretically, analytical or simulated PCF models can be generated by defining the form of cluster size and separation distributions (e.g. exponential or Gaussian). These models can then be fit to experimental curves to extract properties such as the average number of molecules per cluster and the mean cluster size [40**,46]. However, prior knowledge of how clusters are distributed is often not clear, making interpretation of PCF curves highly model-dependent.

Another method, Ripley's K function, $K(r)$, describes the average number of molecules that exist near another molecule within different radius (r) [34]. Conceptually it can be regarded as the cumulative form of the PCF. Similar to the PCF, Ripley's K function can easily discriminate clustered distributions from uniform distributions. In a uniform distribution, $K(r)$, the number of molecules within a radius of r , is proportional to πr^2 . Normalization of $K(r)$ gives rise to Ripley's H function ($H(r) = \sqrt{K(r)/\pi} - r$), which is zero across all distances for a uniform distribution (Figure 3Cii, black curve). For a clustered distribution, $H(r)$ peaks at a characteristic r , which is related to the cluster size; the peak height depends on the mean molecule density in clusters (Figure 3Cii, blue and purple curves). Lagache et al. suggested another normalization of the Ripley's function for quantitative model fitting and easier parameter extracting [47].

Both the PCF and Ripley's functions have been widely used to analyze different types of protein clusters such as membrane receptors [37,40**,46,48]. This is because both functions' shape and amplitude can be intuitively compared with that expected from theoretical calculations to discriminate different models [40**,46]. However, it is important to account for the self-clustering effect due to fluorophore blinking to avoid misinterpretation of clustering at short distances (Section 3.1). In addition, both methods work best when there is only one type of cluster — heterogeneous distributions of multiple cluster types in the same image result in PCF and Ripley's H curves that are difficult to interpret (Figure 3Ci and ii, purple curves). Additionally, quantitative parameter extraction from either function is highly model-dependent, and different cluster properties can generate similar curves with both methods. Thus a prior biological understanding of cluster properties is often required [48,49]. Finally, it is important to note that these functions exhibit edge effects when the neighbor-search region defined by r extends beyond the image boundaries. This results in

significant underestimation of clustering and correlation at large r values. Thus, PCF or Ripley's functions should not be interpreted at r values greater than $\sim 1/3$ of the smallest image dimension without first applying edge-correction methods [50,51].

Detection and quantification of individual clusters using threshold-based methods:

Threshold-based methods are model-independent and rely on segmentation of images or grouping of spots from coordinate lists to isolate individual clusters. When clusters are sparsely distributed, molecules within the same cluster can be grouped together by their proximity using a distance threshold similar to that used for blinking correction (Section 3.1) [43,52]. When clusters are densely-distributed and/or when their sizes have wide distributions, more sophisticated methods are required. For example, tree-clustering methods based on informational theory can be used to group molecules belonging to the same cluster [53]. Cluster boundaries can also be determined using thresholds of fluorescence intensity [32] or the local density of each molecule (e.g. number of nearby molecules or nearest neighbor distance) (Figure 3B) [15,43,54–57]. These thresholds can be normalized (e.g. local density normalized by the average local density) for application to samples with different expression levels or sampling, and are often optimized using simulations that mimic the experimental data [54].

Once cluster boundaries are identified, properties such as cluster area, molecule density, and symmetry can be characterized to describe how different biological states affect these features [32] (Figure 3B, iii–iv). These measurements can reveal assembly mechanisms [56*], differences in oligomeric states [38**,43] or relative constituent stoichiometry [57] in different biological conditions. Ligand binding affinities have also been directly measured in individual cells by counting the number of sparsely-distributed receptor-ligand complexes at different concentrations of fluorescent ligand [58].

3.3 Colocalization—Multicolor superresolution imaging provides the ability to quantify colocalization with spatial resolutions that approach molecule dimensions. Colocalization observed at this scale is more likely to indicate true molecular interactions than colocalization detected with conventional fluorescence microscopy. This precision also results in higher sensitivity to chromatic aberrations between different fluorescence channels. These aberrations can be computationally corrected if first characterized using multi-colored fluorescent beads [59]. Another important consideration is that low sampling frequency caused by inefficient labeling and detection of fluorophores can decrease the apparent colocalization, leading to false negatives. Control experiments for measuring labeling and detection efficiency should be introduced to verify the colocalization results [39,44,45].

Colocalization between different species can often be visually identified as intensity overlap in superresolution images (Figure 4Aiii) [60–62]. Comparing colocalization under different conditions, however, requires statistical analysis of the fluorescence intensity or molecule coordinates of each species. Intensity-based metrics used in conventional fluorescence microscopy have been critically reviewed and compiled into an ImageJ plug-in by Bolte and Cordelieres [63], and can be easily applied to superresolution images [62,64]. For example, the Manders' coefficient calculates the correlation between pixel intensities recorded in two

different channels (Figure 4B, magenta bars). This metric is robust to regions of different sizes used for analysis (Figure 4B i vs. ii), but can produce false-negatives with under-sampled images (Figure 4B iii vs. iv).

Coordinate-based methods analyze the distances between molecules to determine whether two species localize near each other [65,66]. Malkush et al. developed a coordinate-based colocalization (CBC) algorithm to assign a colocalization value for each molecule using the ranked correlation between the number of 'self' vs. 'other' molecules located within varying r values [65]. In addition to allowing comparison of average colocalization values of different images (Figure 4B), this method can provide a histogram or colocalization map, in which areas of high colocalization between two species can be visually identified [65]. We note that this method can exhibit edge effects for the same reasons outlined for the Ripley's and PCF functions above. Because edge effects influence both 'self' and 'other' molecules similarly, the two species can seem highly correlated at large r values, resulting in artificially high CBC values for small region sizes. Thus, it is a good practice to limit the maximum r value (R_{max}) used for the CBC calculations to 1/2 or less of the maximum observed distance between any two molecules (Figure 4B), and it may be useful to adapt an edge-correction method from Ripley's analysis [50,51]. Additionally, region size itself should also be considered when interpreting colocalization values provided by this and other coordinate-based methods. For example, although the two species of clusters in Figure 4Ai are not coincident, they exhibit high colocalization values (Figure 4Bi) because both species are localized to the same general sub-region (Figure 4Bii). When only the sub-region is analyzed, low colocalization values are observed (Figure 4Bii). Hence, region size affects the interpretation of the colocalization values, and direct comparison between different images requires comparable region selection.

Other commonly-used coordinate-based methods are the aforementioned Ripley's functions and PCF (Section 3.2). Both methods can be derived to accommodate multiple species and thus describe the degree of colocalization [67–69]. These methods generate distance-dependent cross-correlation curves of the whole image rather than providing colocalization values for individual molecules (Figure 4C). Although the amplitudes of these cross-correlation curves depend on the size of the region of interest (Figure 4C i vs. ii), the distance value at which maximal correlation values are observed, or the average correlation decay length, is absolute and indicative of the average colocalization displacement between the two species [40**]. Cross-correlation analysis is also robust to under-sampling, which adds noise but does not alter curve shapes (Figure 4C iii vs. iv). Moreover, cross-correlation analysis can distinguish random colocalization that results from high molecule densities from co-clustering (Figure 4C iii vs v). Note that these cross-correlation-based methods are not affected by the overcounting artifacts caused by fluorophore blinking, as they all calculate the correlation of one species with the other but not with itself [40**]. This feature could be utilized to analyze whether a single species self-clusters with itself by labeling the same species with two different labels.

Outlook

Methods to analyze superresolution images continue to evolve, but have already provided invaluable insight into the structures and functions of biological complexes *in vivo*. We hope that the presented examples emphasize the utility of single-molecule localization-based superresolution imaging, and also the caution that should be taken to choose the best-suited analysis methods and controls for each application. The methods presented here often rely on experimental validation using critical control samples, which are becoming more accessible through the development of robust toolkits to generate reliable control constructs [70].

Many of the methods presented here are limited to two dimensions, but we expect rapid adaptation of these tools for the growing number of 3D implementations of PALM and STORM (see Klein et al. for review) [71]. In addition, new tools to analyze the dynamic live-cell data generated by faster localization methods (Section 1), perhaps by adapting single-molecule tracking methodologies [72], are expected to develop. Finally, we anticipate that the next revolution in superresolution imaging will involve the ability to simultaneously determine the activity of single molecules in addition to their locations, perhaps by incorporating other fluorescence techniques such as FRET [73].

Acknowledgments

This work was supported by grants from the National Institutes of Health (1R01GM086447-01A2) and the National Science Foundation (EAGER MCB1019000).

References

Highlighted references

1. Abbe E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*. 1873; 9:413–418.
2. Hell SW. Microscopy and its focal switch. *Nat Meth*. 2009; 6:24–32.
3. Curthoys NM, Mlodzianoski MJ, Kim D, Hess ST. Simultaneous multicolor imaging of biological structures with fluorescence photoactivation localization microscopy. *J Vis Exp*. 2013:e50680. [PubMed: 24378721]
4. Dempsey GT. A user's guide to localization-based super-resolution fluorescence imaging. *Methods Cell Biol*. 2013; 114:561–592. [PubMed: 23931523]
5. Sengupta P, van Engelenburg SB, Lippincott-Schwartz J. Superresolution imaging of biological systems using photoactivated localization microscopy. *Chem Rev*. 2014; 114:3189–3202. [PubMed: 24417572]
6. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*. 2006; 313:1642–1645. [PubMed: 16902090]
7. Rust MJ, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods*. 2006; 3:793–795. [PubMed: 16896339]
8. Heilemann M, van de Linde S, Schüttelpelz M, Kasper R, Seefeldt B, Mukherjee A, Tinnefeld P, Sauer M. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl*. 2008; 47:6172–6176. [PubMed: 18646237]
9. Hess ST, Girirajan TP, Mason MD. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys J*. 2006; 91:4258–4272. [PubMed: 16980368]

10. Shroff H, Galbraith CG, Galbraith JA, Betzig E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat Methods*. 2008; 5:417–423. [PubMed: 18408726]
11. Thompson RE, Larson DR, Webb WW. Precise nanometer localization analysis for individual fluorescent probes. *Biophys J*. 2002; 82:2775–2783. [PubMed: 11964263]
12. Mortensen KI, Churchman LS, Spudich JA, Flyvbjerg H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat Methods*. 2010; 7:377–381. [PubMed: 20364147]
13. Bates M, Huang B, Dempsey GT, Zhuang X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*. 2007; 317:1749–1753. [PubMed: 17702910]
14. Coltharp C, Kessler RP, Xiao J. Accurate construction of photoactivated localization microscopy (PALM) images for quantitative measurements. *PLoS One*. 2012; 7:e51725. [PubMed: 23251611]
15. Tarancon Diez L, Bonsch C, Malkusch S, Truan Z, Munteanu M, Heilemann M, Hartley O, Endesfelder U, Furstenberg A. Coordinate-based co-localization-mediated analysis of arrestin clustering upon stimulation of the C-C chemokine receptor 5 with RANTES/CCL5 analogues. *Histochem Cell Biol*. 2014
16. Shannon EC. Communication in the Presence of Noise. *Proceedings of the IEEE*. 1949; 86:447–457.
17. Jones SA, Shim S-H, He J, Zhuang X. Fast, three-dimensional super-resolution imaging of live cells. *Nat Meth*. 2011; 8:499–505.
18. Nieuwenhuizen RP, Lidke KA, Bates M, Puig DL, Grunwald D, Stallinga S, Rieger B. Measuring image resolution in optical nanoscopy. *Nat Methods*. 2013; 10:557–562. [PubMed: 23624665] The authors discuss the effects of both localization precision and sampling on image resolution. They also provide an correlation-based method borrowed from electron microscopy to quantify the resolution achieved in a superresolution image.
19. Fu G, Huang T, Buss J, Coltharp C, Hensel Z, Xiao J. In vivo structure of the E. coli FtsZ-ring revealed by photoactivated localization microscopy (PALM). *PLoS One*. 2010; 5:e12682. [PubMed: 20856929]
20. Holden SJ, Uphoff S, Kapanidis AN. DAOSTORM: an algorithm for high- density super-resolution microscopy. *Nat Methods*. 2011; 8:279–280. [PubMed: 21451515]
21. Wang Y, Quan T, Zeng S, Huang ZL. PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy. *Opt Express*. 2012; 20:16039–16049. [PubMed: 22772294]
22. Cox S, Rosten E, Monypenny J, Jovanovic-Talisman T, Burnette DT, Lippincott-Schwartz J, Jones GE, Heintzmann R. Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nat Methods*. 2012; 9:195–200. [PubMed: 22138825]
23. Zhu L, Zhang W, Elnatan D, Huang B. Faster STORM using compressed sensing. *Nat Methods*. 2012; 9:721–723. [PubMed: 22522657]
24. Mukamel EA, Babcock H, Zhuang X. Statistical deconvolution for superresolution fluorescence microscopy. *Biophys J*. 2012; 102:2391–2400. [PubMed: 22677393]
25. Deng Y, Sun M, Lin PH, Ma J, Shaevitz JW. Spatial Covariance Reconstructive (SCORE) Super-Resolution Fluorescence Microscopy. *PLoS One*. 2014; 9:e94807. [PubMed: 24788039]
26. Ober RJ, Ram S, Ward ES. Localization accuracy in single-molecule microscopy. *Biophys J*. 2004; 86:1185–1200. [PubMed: 14747353]
27. Zhang Y, Gu L, Chang H, Ji W, Chen Y, Zhang M, Yang L, Liu B, Chen L, Xu T. Ultrafast, accurate, and robust localization of anisotropic dipoles. *Protein Cell*. 2013; 4:598–606. [PubMed: 23744341]
28. Small A, Stahlheber S. Fluorophore localization algorithms for super-resolution microscopy. *Nat Methods*. 2014; 11:267–279. [PubMed: 24577277] This review comprehensively introduces the principles of localization and compares the precision and computational speed of different algorithms. The authors also summarize methods to detect overlapping molecules.
29. Small AR, Parthasarathy R. Superresolution localization methods. *Annu Rev Phys Chem*. 2014; 65:107–125. [PubMed: 24274701]

30. Hoogendoorn E, Crosby KC, Leyton-Puig D, Breedijk RM, Jalink K, Gadella TW, Postma M. The fidelity of stochastic single-molecule super-resolution reconstructions critically depends upon robust background estimation. *Sci Rep*. 2014; 4:3854. [PubMed: 24458236]
31. Henriques R, Lelek M, Fornasiero EF, Valtorta F, Zimmer C, Mhlanga MM. QuickPALM: 3D real-time photoactivation nanoscopy image processing in Image. *J. Nat Methods*. 2010; 7:339–340.
32. Buss J, Coltharp C, Huang T, Pohlmeier C, Wang SC, Hatem C, Xiao J. In vivo organization of the FtsZ-ring by ZapA and ZapB revealed by quantitative super-resolution microscopy. *Mol Microbiol*. 2013; 89:1099–1120. [PubMed: 23859153]
33. Baddeley D, Cannell MB, Soeller C. Visualization of localization microscopy data. *Microsc Microanal*. 2010; 16:64–72. [PubMed: 20082730]
34. Penttinen A, Stoyan D. Recent applications of point process methods in forestry statistics. *Statistical Science*. 2000; 15:61–78.
35. Annibale P, Vanni S, Scarselli M, Rothlisberger U, Radenovic A. Quantitative photo activated localization microscopy: unraveling the effects of photoblinking. *PLoS One*. 2011; 6:e22678. [PubMed: 21818365]
36. Sengupta P, Lippincott-Schwartz J. Quantitative analysis of photoactivated localization microscopy (PALM) datasets using pair-correlation analysis. *Bioessays*. 2012; 34:396–405. [PubMed: 22447653]
37. Lee SH, Shin JY, Lee A, Bustamante C. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc Natl Acad Sci U S A*. 2012; 109:17436–17441. [PubMed: 23045631]
38. Puchner EM, Walter JM, Kasper R, Huang B, Lim WA. Counting molecules in single organelles with superresolution microscopy allows tracking of the endosome maturation trajectory. *Proc Natl Acad Sci U S A*. 2013; 110:16015–16020. [PubMed: 24043832] This study introduces many crucial control constructs and analyses for molecule counting applications.
39. Zhao ZW, Roy R, Gebhardt JC, Suter DM, Chapman AR, Xie XS. Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. *Proc Natl Acad Sci U S A*. 2014; 111:681–686. [PubMed: 24379392]
40. Veatch SL, Machta BB, Shelby SA, Chiang EN, Holowka DA, Baird BA. Correlation functions quantify superresolution images and estimate apparent clustering due to over-counting. *PLoS One*. 2012; 7:e31457. [PubMed: 22384026] This is a thorough discussion of pair correlation function principles and applications. Understanding these concepts is crucial to successful utilizing this method.
41. Wang S, Moffitt JR, Dempsey GT, Xie XS, Zhuang X. Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging. *Proceedings of the National Academy of Sciences*. 2014
42. Durisic N, Laparra-Cuervo L, Sandoval-Alvarez A, Borbely JS, Lakadamyali M. Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. *Nat Meth*. 2014; 11:156–162. The authors discuss the importance of assessing the detection efficiency of photactivatable fluorescent proteins, and provide a thorough method of quantifying efficiency using a structure of known stoichiometry.
43. Nan X, Collisson EA, Lewis S, Huang J, Tamguney TM, Liphardt JT, McCormick F, Gray JW, Chu S. Single-molecule superresolution imaging allows quantitative analysis of RAF multimer formation and signaling. *Proc Natl Acad Sci U S A*. 2013; 110:18519–18524. [PubMed: 24158481]
44. Annibale P, Scarselli M, Greco M, Radenovic A. Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy. *Optical Nanoscopy*. 2012; 1:9.
45. Renz M, Daniels BR, Vámosi G, Arias IM, Lippincott-Schwartz J. Plasticity of the asialoglycoprotein receptor deciphered by ensemble FRET imaging and single-molecule counting PALM imaging. *Proceedings of the National Academy of Sciences*. 2012; 109:E2989–E2997.

46. Sengupta P, Jovanovic-Talisman T, Skoko D, Renz M, Veatch SL, Lippincott-Schwartz J. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat Methods*. 2011; 8:969–975. [PubMed: 21926998]
47. Lagache T, Lang G, Sauvonnnet N, Olivo-Marin J-C. Analysis of the Spatial Organization of Molecules with Robust Statistics. *PLoS One*. 2013; 8:e80914. [PubMed: 24349021]
48. Kiskowski MA, Hancock JF, Kenworthy AK. On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophys J*. 2009; 97:1095–1103. [PubMed: 19686657]
49. Hess ST, Gould TJ, Gudheti MV, Maas SA, Mills KD, Zimmerberg J. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *Proc Natl Acad Sci U S A*. 2007; 104:17370–17375. [PubMed: 17959773]
50. Haase P. Spatial pattern analysis in ecology based on Ripley's K-function: Introduction and methods of edge correction. *Journal of Vegetation Science*. 1995; 6:575–582.
51. Ripley, BD. *Statistical Inference for Spatial Processes*. New York, NY, USA: Cambridge University Press; 1989.
52. Gunzenhauser J, Olivier N, Pengo T, Manley S. Quantitative super-resolution imaging reveals protein stoichiometry and nanoscale morphology of assembling HIV-Gag virions. *Nano Lett*. 2012; 12:4705–4710. [PubMed: 22906128]
53. Greenfield D, McEvoy AL, Shroff H, Crooks GE, Wingreen NS, Betzig E, Liphardt J. Self-organization of the Escherichia coli chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol*. 2009; 7:e1000137. [PubMed: 19547746]
54. MacGillavry HD, Song Y, Raghavachari S, Blanpied TA. Nanoscale scaffolding domains within the postsynaptic density concentrate synaptic AMPA receptors. *Neuron*. 2013; 78:615–622. [PubMed: 23719161]
55. Scarselli M, Annibale P, Radenovic A. Cell type-specific beta2-adrenergic receptor clusters identified using photoactivated localization microscopy are not lipid raft related, but depend on actin cytoskeleton integrity. *J Biol Chem*. 2012; 287:16768–16780. [PubMed: 22442147]
56. Bar-On D, Wolter S, van de Linde S, Heilemann M, Nudelman G, Nachliel E, Gutman M, Sauer M, Ashery U. Super-resolution imaging reveals the internal architecture of nano-sized syntaxin clusters. *Journal of Biological Chemistry*. 2012 This study highlights the application-specific nature of cluster segmentation. The authors present several validations of their multistep method that utilizes both image-processing and local-density methods to identify and segment protein clusters.
57. Ori A, Banterle N, Iskar M, Andres-Pons A, Escher C, Khanh Bui H, Sparks L, Solis-Mezarino V, Rinner O, Bork P, et al. Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol Syst Biol*. 2013; 9:648. [PubMed: 23511206]
58. Dietz MS, Fricke F, Kruger CL, Niemann HH, Heilemann M. Receptor-ligand interactions: binding affinities studied by single-molecule and super-resolution microscopy on intact cells. *Chemphyschem*. 2014; 15:671–676. [PubMed: 24772464]
59. Fletcher PA, Scriven DR, Schulson MN, Moore ED. Multi-image colocalization and its statistical significance. *Biophys J*. 2010; 99:1996–2005. [PubMed: 20858446]
60. Ribeiro SA, Vagnarelli P, Dong Y, Hori T, McEwen BF, Fukagawa T, Flors C, Earnshaw WC. A super-resolution map of the vertebrate kinetochore. *Proc Natl Acad Sci U S A*. 2010; 107:10484–10489. [PubMed: 20483991]
61. Lew MD, Lee SF, Ptacin JL, Lee MK, Twieg RJ, Shapiro L, Moerner WE. Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live *Caulobacter crescentus*. *Proc Natl Acad Sci U S A*. 2011; 108:E1102–E1110. [PubMed: 22031697]
62. Zhao J, Bruck S, Cemurski S, Zhang L, Butler B, Dani A, Cooper JA, Shaw AS. CD2AP links cortactin and capping protein at the cell periphery to facilitate formation of lamellipodia. *Mol Cell Biol*. 2013; 33:38–47. [PubMed: 23090967]
63. Bolte S, Cordelieres FP. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc*. 2006; 224:213–232. [PubMed: 17210054]

64. Bielopolski N, Lam AD, Bar-On D, Sauer M, Stuenkel EL, Ashery U. Differential interaction of tomosyn with syntaxin and SNAP25 depends on domains in the WD40 beta-propeller core and determines its inhibitory activity. *J Biol Chem*. 2014
65. Malkusch S, Endesfelder U, Mondry J, Gelleri M, Verveer PJ, Heilemann M. Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem Cell Biol*. 2012; 137:1–10. [PubMed: 22086768]
66. Rossy J, Cohen E, Gaus K, Owen DM. Method for co-cluster analysis in multichannel single-molecule localisation data. *Histochem Cell Biol*. 2014
67. Sherman E, Barr V, Manley S, Patterson G, Balagopalan L, Akpan I, Regan CK, Merrill RK, Sommers CL, Lippincott-Schwartz J, et al. Functional nanoscale organization of signaling molecules downstream of the T cell antigen receptor. *Immunity*. 2011; 35:705–720. [PubMed: 22055681]
68. Lehmann M, Rocha S, Mangeat B, Blanchet F, Uji IH, Hofkens J, Piguet V. Quantitative multicolor superresolution microscopy reveals tetherin HIV-1 interaction. *PLoS Pathog*. 2011; 7:e1002456. [PubMed: 22194693]
69. Notelaers K, Rocha S, Paesen R, Swinnen N, Vangindertael J, Meier JC, Rigo JM, Ameloot M, Hofkens J. Membrane distribution of the glycine receptor alpha3 studied by optical super-resolution microscopy. *Histochem Cell Biol*. 2014
70. Schmied JJ, Raab M, Forthmann C, Pibiri E, Wünsch B, Dammeyer T, Tinnefeld P. DNA origami-based standards for quantitative fluorescence microscopy. *Nat. Protocols*. 2014; 9:1367–1391.
71. Klein T, Proppert S, Sauer M. Eight years of single-molecule localization microscopy. *Histochem Cell Biol*. 2014
72. Manley S, Gillette JM, Patterson GH, Shroff H, Hess HF, Betzig E, Lippincott-Schwartz J. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat Methods*. 2008; 5:155–157. [PubMed: 18193054]
73. Grecco HE, Verveer PJ. FRET in cell biology: still shining in the age of super-resolution? *Chemphyschem*. 2011; 12:484–490. [PubMed: 21344589]

- Single-molecule superresolution images enable molecule counting and pattern identification.
- Spatial resolution depends on both localization precision and sampling frequency.
- Under- and over-counting artifacts can be corrected by using proper controls.
- Molecule distributions can be characterized by correlation- or threshold-based analyses.
- Colocalization can be characterized by intensity- or coordinate-based analyses.

Figure 1. Spatial resolution of single-molecule super-resolution image

(A) Schematic of improved spatial resolution (red) in comparison with diffraction-limited single molecule PSF (blue). Blue and red shaded regions represent the standard deviations of the PSF (σ) and localization precision (σ_{loc}), corresponding to the provided equations. (λ : fluorescence wavelength; NA : numeric aperture; N : detected photon number; a : pixel size; b : background noise) [11]. (B) Effect of photon number and sampling frequency on spatial resolution. Higher photon numbers improve the localization precision, σ_{loc} (black dashed curve). Blue and red dashed lines indicate the Nyquist limitation, σ_{Nyq} , under different sampling frequencies. The improvement of overall image resolution (blue and red solid curves) is limited by both localization precision and sampling frequency. (C) Experimental measurement of localization precision. Left: Multiple detections of a single molecule (e.g. organic fluorophores, red spots). Right: Molecules detected only a few times (e.g. fluorescent proteins). R is the distance between localizations in continuous frames. Blue circles indicate experimental localization precision (σ_{loc}^{exp}). Fitting the histograms (gray) with the equations shown in each plot can estimate the precision. (D) Effect of localization precision and sampling frequency on superresolution image reconstruction. The two green lines represent two 100 nm long and 1 nm wide parallel structures separated by 60 nm. Left column: Simulated localizations (red) with $\sigma_{loc} = 11\text{ nm}$ or 22 nm and sampling frequency at 500 or 50,000 μm^{-2} . Right column: Superresolution image constructed as in Figure 2B. Low

precision (i) or sampling (ii) obscures the structures. Sufficiently high precision and sampling frequency are required to distinguish and recover the structures (iii).

Figure 2. Super-resolution image generation and dimension measurement

(A) Simulated clusters (70 nm FWHM) displayed as a diffraction-limited image (i), superresolution Gaussian plot (ii), and superresolution 2D histogram (iii). Sigma for each Gaussian spot in (ii) is specified as 15 nm localization precision. (B) Underlying structural cross section profiles (green) are convolved with Gaussian-distributed localization precision (magenta) to generate apparent profiles in superresolution images (gray). Fitting the apparent profile with a Gaussian distribution (cyan) yields the apparent width, $FWHM_{app}$. The true FWHM can be inferred using the relationship

$FWHM_{true} = \sqrt{FWHM_{app}^2 + FWHM_{loc}^2}$, which is only valid for inherently Gaussian-distributed structures (iii and iv), and yields erroneous measurements for other structures (i and ii). Prior knowledge of the underlying structures can yield the correct width via deconvolution ($width_{decon}$ and $FWHM_{decon}$). (C) Cluster FWHM measured from the 2D histogram (red) and Gaussian plot (blue) plotted against true FWHM. Insets: representative measurements of 70 nm clusters (white dotted outlines in A). Y-intercepts represent smallest measurable FWHM by each method, which is determined by localization precision (magenta).

Figure 3. Analysis of molecule clustering

(A) Simulated superresolution images of molecules distributed randomly (i), in small clusters (ii), in large clusters (iii), and in clusters of multiple sizes (iv). (B) Cluster identification and analysis from local density maps. (i) Local density of molecules is represented by the number of molecules within $100 \text{ nm} \times 100 \text{ nm}$ pixel. (ii) Clusters are clearly distinguished after removing molecules below a local density threshold (in this example, the threshold is the mean local density value). (iii) Clusters are segmented from a binary image generated from ii. The number of constituent molecules are shown for two clusters. (iv) Coloring each cluster by its area illustrates three populations of cluster sizes: large (white), medium (orange), and small (red). (C) Statistical analyses by pair-correlation function (PCF) (i) and Ripley's functions (ii) show distinct curve shapes for the simulations shown in A. For homogenous cluster distributions (blue, purple), cluster sizes can be approximated from the pair-correlation decay length (i) or the peak position of Ripley's H function (ii). However, extracting the mixed cluster sizes (red) is not intuitive for either curve. Scale bars, 500 nm.

Figure 4. Colocalization analysis

(Ai to Avi) Representative superresolution images with two simulated species (green and purple; white represents overlap). (i and ii) The two species form self clusters and are randomly distributed with respect to each other in a large region (i) or a small region (ii). (iii and iv) The two species form self clusters and colocalize with each other, but are detected with high (iii) and low (iv) efficiency respectively. (v and vi) the two species do not form self clusters, are randomly distributed, and detected with high (v) and low (vi) efficiency. (B). Comparison of the coordinate-based colocalization (CBC) algorithm [65] and Manders' Coefficient analysis of images Ai to Avi. CBC values were calculated using $R_{max} = 1/3$ of the maximum distance between any two molecules. (C) Comparison of cross-correlation analysis of images Ai to Avi.