Image scanning microscopy: an overview

E. N. WARD & R. PAL
Department of Chemistry, Durham University, Durham, UK

Key words. Image scanning microscopy, structured illumination, super-resolution microscopy.

Summary
For almost a century, the resolution of optical microscopy was thought to be limited by Abbé’s law describing the diffraction limit of light. At the turn of the millennium, aided by new technologies and fluorophores, the field of optical microscopy finally surpassed the diffraction barrier: a milestone achievement that has been recognized by the 2014 Nobel Prize in Chemistry. Many super-resolution methods rely on the unique photophysical properties of the fluorophores to improve resolution, posing significant limitations on biological imaging, such as multicoloured staining, live-cell imaging and imaging thick specimens. Structured Illumination Microscopy (SIM) is one branch of super-resolution microscopy that requires no such special properties of the applied fluorophores, making it more versatile than other techniques. Since its introduction in biological imaging, SIM has proven to be a popular tool in the biologist’s arsenal for following biological interaction and probing structures of nanometre scale. SIM continues to see much advancement in design and implementation, including the development of Image Scanning Microscopy (ISM), which uses patterned excitation via either predefined arrays or raster-scanned single point-spread functions (PSF). This review aims to give a brief overview of the SIM and ISM processes and subsequent developments in the image reconstruction process. Drawing from this, and incorporating more recent achievements in light shaping (i.e. pattern scanning and super-resolution beam shaping), this study also intends to suggest potential future directions for this ever-expanding field.

Background
Fluorescence microscopy stands out as possibly the most ubiquitous tool in biological imaging. However, this versatile technique suffers from a critical barrier in resolution, which limits its potential to study samples at the nanometre scale. First formalized in 1873 by Ernst Abbé, this resolution barrier is dependent on the wavelength of light and the numerical aperture of the imaging system; itself a function of the refractive index of the imaging media and the angle of light that can be received from the focal plane of the objective lens (Abbé, 1873). For visible light and a high numerical aperture objective, this typically limits resolution to ~200 nm in the lateral plane and ~700 nm in the axial direction. The laser scanning confocal microscope (LSCM) was the first microscope to reach this resolution barrier and has become the most widespread imaging tool in biological imaging. The key advantage of the LSCM is a pinhole in the detection arm of the optical path to reject out-of-focus light, allowing structures to be studied in 3D by imaging only a single focal plane.

Super-resolution microscopy (nanoscopy) is a more recent development in the field of fluorescence microscopy, and describes any imaging technique capable of breaking Abbé’s diffraction limit. To date, there has been a vast number of nanoscopic techniques conceived, and innumerable examples of new scientific discoveries resulting from these techniques. So far, the highest resolution methods, such as the single-molecule localization microscopy techniques – Photo-Activation Localization Microscopy (PALM) (Betzig et al., 2006) and Stochastic Optical Reconstruction Microscopy (STORM) (Rust et al., 2006) are all exploiting the properties of the applied fluorophores. Stimulated emission depletion (STED) microscopy also harnesses the unique properties of fluorophores combined with a doughnut-shaped depletion beam that travels along the same path as a classical excitation beam. The generated depletion beam acts to reduce the effective size of the excitation spot (Hell & Wichmann, 1994).

Structured illumination microscopy (SIM) describes any subdiffraction technique that involves the use of patterned excitation. (Although structured illumination imaging is possible without fluorescent response, this review will be limited to its application in fluorescence microscopy.) The first examples of SIM involved the use of opposing objectives focused onto the same focal plane (Hell et al., 1994; Gustafsson et al., 1995). In this configuration, it is possible to generate standing waves or interference patterns in the axial plane with periods shorter than the axial resolution of the objectives. This has the effect of reducing the thickness of the plane illuminated...
and subsequently increasing axial resolution. Lateral increase necessitates the use of laterally patterned excitation and was first achieved at the turn of the millennium (Heintzmann & Cremer, 1999; Gustafsson, 2000). Since then, it has taken off as one of the key methods in nanoscopy. The great advantages of SIM are that it is compatible with any currently available fluorophore and that it uses modest excitation powers. This allows existing protocols to be used and facilitates live-cell imaging.

Theory of super-resolution SIM

For any point, \( \tilde{r} \), in an image recorded from a wide-field microscope, the intensity, \( D(\tilde{r}) \), is given by

\[
D(\tilde{r}) = (S(\tilde{r}) \cdot E(\tilde{r})) \otimes \text{PSF}(\tilde{r}).
\]  

(1)

For the correct description of a wide-field microscope, the case of incoherent illumination is assumed throughout (Goodman 2004). \( S(\tilde{r}) \) is the structure or distribution of fluorophores in the sample; \( E(\tilde{r}) \) is the excitation light amplitude and \( \text{PSF}(\tilde{r}) \) is the detection Point-Spread Function (PSF). \( \otimes \) denotes the mathematical convolution operation which has the effect of blending two functions. The PSF of the system can be conceived as the result in the image plane of an infinitesimal emitter in the sample plane. It also represents the smallest volume to which light can be focused. When considering the resolution limit of an imaging system, it is easier to consider the associated frequency space. Taking the Fourier transform of the function converts the intensity distribution from real to frequency space:

\[
\tilde{D}(\tilde{k}) = \tilde{S}(\tilde{k}) \otimes \tilde{E}(\tilde{k}) \otimes \text{OTF}(\tilde{k}).
\]  

(2)

The tildes denote the Fourier transforms of the component functions, and the convolution and product have been swapped according to the mathematical definition of a convolution. The Optical Transfer Function, \( \text{OTF}(k) \), is the direct Fourier transform of the PSF. In frequency space, higher spatial frequencies correspond to a better-resolved image. The OTF of an imaging system acts as low-pass filter, cutting off high spatial frequencies. The general SIM process is depicted in Figure 1. In the simplest case of 2D SIM, a striped sinusoidal pattern is used and the function \( E(\tilde{r}) \) becomes

\[
E(\tilde{r}) = E_0 (1 + \cos(\tilde{k}_0 \tilde{r} + \varphi)).
\]  

(3)

Here, the vector \( \tilde{k}_0 \) describes the frequency and direction of the pattern, and \( \varphi \) is the phase. When substituting this into the Eq. (3), we get

\[
\tilde{D}(\tilde{k}) = E_0 [\tilde{S}(\tilde{k}) + 0.5\tilde{S}(\tilde{k} + \tilde{k}_0)e^{i\varphi} + 0.5\tilde{S}(\tilde{k} - \tilde{k}_0)e^{-i\varphi}].
\]  

(4)

Looking at this, the detected Fourier spectrum is now a linear superposition of three zones of the frequency space of the sample structure. To extract the extra information, it is necessary to move the new frequency components to their correct place in the final image spectrum. This is achieved by capturing three separate images with different illumination pattern phases to isolate each sampled region of frequency space. The extra frequency information can then be assigned to its correct place in frequency space. Since the frequency components are only shifted in the direction of \( \tilde{k}_0 \), it is necessary to repeat this process at three directions of illumination pattern to achieve isotropic resolution increase (Fig. 1). SIM is particularly sensitive to any movement of the sample and to changes in its fluorescent response while these pictures are being acquired. As such, these images must be captured in quick succession and with a suitably low illumination intensity to minimize sample drift, natural cell homeostatic movement and organelle rearrangement as well as eliminating photobleaching.

There have been a number of improvements to this core SIM implementation. Of particular note is Saturated SIM (SSIM) (Gustafsson, 2005) which achieves a theoretically unlimited resolution by using the nonlinear response of fluorophores at high excitation intensities. Above a certain intensity threshold, the fluorescent response saturates. This is seen as a ‘levelling off’ in the sinusoidal to a square pattern. This square pattern introduces higher spatial frequencies into the excitation pattern, allowing for sampling of a greater region of the object’s frequency space. To extract these extra frequencies, more phase shifts are required for each direction of the pattern, extending acquisition time. However, although superior resolution is achievable, the increased imaging time and the higher excitation power are normally incompatible with live-cell imaging. SSIM is also very susceptible to photobleaching, as any difference in response between excitation patterns results in artefacts after image reconstruction.

Image scanning microscopy (ISM)

One of the more recent advancements in SIM has been ISM which was achieved in 2010 by Müller & Enderlein (Müller & Enderlein, 2010), despite having been previously described some years earlier (Sheppard, 1988). The underlying principle of ISM can be understood as extracting the inherent super-resolution information from an LSCM. The origin of this extra information can be conceptualized in two different ways. The first description is based on the idea of the overlap of excitation and emission PSFs in a confocal microscope (Sheppard, 1988; Sheppard et al., 2013); the second sought to describe this as an SIM technique (Müller & Enderlein, 2010). In a scanning microscope, a diffraction-limited spot is raster scanned across the sample and the image is built up in a pixel-by-pixel fashion. This means that, for every point in the scan, the excitation pattern is the excitation PSF. By definition, the Fourier transform of this diffraction-limited spot contains all the spatial frequencies permitted by the
objective. As with all patterned illumination, the spatial frequency components of the sample are mixed with those of the excitation pattern, meaning high spatial frequencies of the sample are moved into the range of the detection OTF. In the case of the point-scanning microscope, the highest frequency component of the excitation pattern is the cut-off frequency of the excitation optical transfer function. This means that the maximum spatial frequency moved into the supported region of the detection OTF is double that which is usually gathered.

In an LSCM, some of this super-resolution information can be collected simply by closing the pinhole. However, in practice, the large amount of light rejected by a smaller pinhole reduces the signal-to-noise ratio to prohibitively low levels. ISM works by recovering this lost super-resolution information. The optical configuration required to achieve ISM on a scanning microscope is relatively simple and involves recording the signal that passes through the pinhole on an array detector and capturing an image at every scan position. The simplest way to recover a final super-resolution image is pixel reassignment, shown in Figure 2 (Sheppard et al., 2013). In practice, pixel reassignment is relatively simple to accomplish. For each scan position, the image acquired is shrunk by some factor, and added to a running-total image, centred at the scan position of the beam. The degree of shrinking is based on several factors, including the Stoke’s shift of the fluorophores used. After a complete picture has been reconstructed, further resolution enhancement can be achieved by Fourier reweighting. It should be noted that this pixel reassignment technique has also achieved all-optically – without the need for complicated postprocessing – by Roth et al. (2013) and more recently with a modified spinning disc microscope (Azuma & Kei, 2015) for faster acquisition.

Developments of ISM

From these early steps, ISM has undergone significant improvement to be transformed into a more robust microscopic technique. The first hurdle to overcome was the extended acquisition time associated with all scanning methods, that is even further prolonged in ISM. Since the light gathered is spread out over the imaging chip of a camera, the signal recorded on the array detector is weaker, and subsequently longer pixel dwell times are required. This means there is a total image acquisition time of at least 25 s for a modest 4 μm × 4 μm field of view. This can be greatly reduced if more than one excitation spot is used simultaneously and the whole field of view is recorded, analogous with spinning disc confocal microscopy. This was first accomplished by York et al., who used a Digital Micromirror Device (DMD) to illuminate the sample in a process they termed Multifocal Structured Illumination Microscopy (MSIM) (Fig. 3) (York et al., 2012). Using the DMD, they reported speeds of up to 1 Hz for a 50 μm × 50 μm field of view: a significant improvement on the scanning method, now offering a temporal resolution suitable for basic live-cell imaging. Generating patterns for MSIM with DMDs, wastes most of the available excitation light due to the limited number of active pixels used in the array. Holography offers one possible solution to overcome this issue (Jesacher & Ritsch-Marte, 2016) along with the possibility to generate a wider range of more complicated structures. So far waveform shaping with holography in SIM and MSIM has not yet been accomplished but in the future may allow for more sophisticated pattern projection. Furthermore, since MSIM employs wide-field detection, the optical sectioning capability is lost. This can, in part, be recovered through the process of ‘digital pinholing’. To achieve this, the location of the excitation spots in each of the raw images is determined and a mask is then applied around this point, rejecting light gathered on surrounding pixels. This has the effect of partially removing light from outside the focal plane. An alternative approach to retaining optical sectioning while using multiple excitation spots has also been demonstrated by Schulz et al., who combined a spinning-disc microscope with a microsecond pulsed laser (Schulz et al., 2013).

As well as reducing total image acquisition times, lately there have been substantial advances in the image
reconstruction process. Newer reconstruction procedures for ISM have focused on using maximum-likelihood deconvolution (MLD) algorithms. Put simply, MLD is an iterative process that maximizes the probability that an estimated sample structure will generate the images acquired under the illumination patterns used. The classical diffraction-limited image is taken as the initial estimate of the sample. The computer then predicts the fluorescent response of the estimated sample to each illumination pattern. From the differences between the acquired and predicted images, an update step is calculated and applied to the initial estimate to generate a new estimate. The process is then repeated either until the update step reaches a predefined minimum, or until a user-defined iteration limit is reached. The most popular MLD algorithms are joint Richardson–Lucy deconvolution (Ingaramo et al., 2014; Ströhl & Kaminski, 2015) and pattern-illuminated Fourier Ptychography (Dong et al., 2014). The difference between these is the way in which the update step is calculated, and the relative strengths of these methods and the mathematical background has recently been reviewed extensively by Chakrova et al. (2016). MLD has been shown to outperform pixel reassignment in terms of both resolution improvement and signal-to-noise ratio. This is in part because MLD can be adapted to account for one or more of the different types of noise associated with image acquisition, specifically: Poisson noise originating from the low photon counts; and Gaussian noise originating from the readout noise of the camera. This is in contrast to pixel reassignment where, rather than being suppressed, noise is amplified during image reconstruction. MLDs also eliminate the need to apply Fourier reweighting onto the image after reconstruction and, under certain circumstances, improve resolution beyond ISM alone by estimating super-resolution information from predefined knowledge of the sample (Heintzmann, 2007). A final key advantage of MLDs is that they are able to extract
structured illumination information in situations where the direct reconstruction process in frequency space is not defined. In fact, MLD can even be implemented in cases where the illumination pattern is unknown (Mudry et al., 2012), though these methods are generally outperformed by using known illumination patterns (Chakrova & Rieger, 2015). Pixel reassignment, however, is based on the underlying assumption that the emission and detection PSFs are scaled versions of each other, and as such is resultantly applicable in very specific circumstances.

Limitations and drawbacks

As with all techniques, ISM suffers from some limitations. These are particularly apparent when compared to other nanoscopic methods (Table 1). Since the structured illumination information comes from a diffraction-limited pattern, it is limited to only a 2-fold lateral resolution increase; this poses no overall improvement over simpler forms of SIM which use a striped pattern. Furthermore, scanning ISM requires much greater acquisition time than LSCM or SIM. Multispot ISM offers an improvement to temporal resolution but, unless it is operated on a spinning disc microscope, it loses the confocality offered by scanning systems. However, total internal reflection excitation and axially patterned excitation (Gustafsson et al., 2008) have allowed other forms of SIM to avoid this issue and achieve super-resolution in all three dimensions, something that has not yet been exploited with ISM.

Precise knowledge of the excitation pattern is also imperative to MSIM in both applying digital pinholes and in image reconstruction. The effects of this are particularly apparent when imaging highly scattering or thick tissue samples, where aberrations and noise produce reconstruction artefacts in the final image. This has proven problematic in practical applications of ISM, where the pattern must either be determined by regularly calibrating the system using a test slide (York et al., 2012; Schulz et al., 2013) or by computationally identifying the excitation pattern postacquisition (McGregor et al., 2015; Ströhl & Kaminski, 2015). Both of these techniques have their limitations: calibrating the system is a time-consuming step, and is ineffective if the sample has greatly different optical properties to the test slide. Determining the pattern postacquisition adds further steps to an already computationally intensive technique and, depending on the algorithm used, reconstruction can break down in patterned or sparsely fluorescing samples (McGregor et al., 2015). Combining MSIM with two-photon excitation has been demonstrated to reduce the effects of scattering samples and may offer one solution to this issue (Wawrzusin et al., 2014).

Future directions of ISM

Despite these shortfalls, ISM is an ever-developing method and shows great potentials as a super-resolution technique. A significant limitation to ISM is the relatively small increase in resolution it affords. Since the resolution increase in ISM is achieved by the frequency components of the focal spot, introducing higher-than-classically-allowed spatial frequencies into the excitation spot would allow for a better than 2-fold gain in resolution. One promising direction is merging the fields of PSF engineering and ISM. PSF engineering involves altering an optical system in such a way as to generate a non-classical excitation or detection PSF (Fang et al., 2015). The most prominent example of PSF engineering is the generation of the doughnut-shaped depletion beam used in 2D STED nanoscopy (Hell & Wichmann, 1994). To create this doughnut shape, a spiral phase is imparted into the beam, generating an intensity minimum at the centre of the focal spot. Crucially, the size of this dark region is not diffraction-limited, and as such, preventing fluorescence outside of this central region results in an effective excitation PSF that is considerably smaller than in LSCM. ISM on a scanning STED microscope was recently demonstrated by Laporte et al. (2014), who used pixel reassignment to give a 1.25x improvement in resolution over stand-alone STED nanoscopy. However, it may be possible to extract structured illumination information from a microscope when using only exclusively a pattern of doughnut-shaped excitation spots. Since the width of the central dark spot is subdiffraction, there are higher spatial frequencies in the doughnut spot than a diffraction-limited Gaussian PSF. Figures 4(G) and (H) show an example of simulated MSIM data when using arrays of doughnut PSFs. The subdiffraction structure of the excitation PSFs has given a clear improvement in resolution over MSIM using a Gaussian spot pattern. This pattern could be generated via cavity-modulation of the excitation laser beam, however this method requires costly custom-built individually modulated laser systems to be employed. Furthermore, the generated laser modes could also be difficult to preserve through the rest of the optical configuration employed in ISM (Mader et al., 2015). Laporte et al. also attempted to build on the Saturated SIM method and take advantage of saturation of fluorophores within a single excitation spot. As in SSIM, saturation of fluorophores leads to an effectively flat-topped PSF, containing high spatial frequencies. Although theoretically possible, this method proved unsuccessful, since the intensity of these high-frequency components was too low to be detectable.

Other efforts to bring PSF engineering into ISM have aimed to improve the axial resolution. RESCH (REFocusing after SCanning using a Helical phase engineering) is one such technique that operates in a similar manner to scanning ISM (Jesacher et al., 2015; Roider et al., 2016). As with ISM, the signal from a point-scanning microscope is collected on a camera in the place of a point detector. However, in RESCH, the phase of the collected light is modulated by an Spatial Light Modulator (SLM) in such a way that axial information is encoded into the image captured on the camera. By sampling different regions of the camera image, the fluorescence at different focal

© 2017 The Authors
Journal of Microscopy © 2017 Royal Microscopical Society, 266, 221–228
Table 1. Comparison of the discussed microscopy techniques highlighting typically reported experimental results achieved in key domains of biological imaging.

<table>
<thead>
<tr>
<th></th>
<th>LSCM</th>
<th>STED</th>
<th>SMLM</th>
<th>2D-SIM</th>
<th>3D-SIM</th>
<th>SSIM</th>
<th>ISM</th>
<th>MSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral resolution</td>
<td>~200</td>
<td>~20</td>
<td>~10–20</td>
<td>~150</td>
<td>~120</td>
<td>~50</td>
<td>~150</td>
<td>~145</td>
</tr>
<tr>
<td>Axial resolution</td>
<td>~600</td>
<td>~40</td>
<td>~10–40</td>
<td>No data</td>
<td>~360</td>
<td>No data</td>
<td>No data</td>
<td>~400</td>
</tr>
<tr>
<td>Frame rate (Hz)</td>
<td>&gt;1</td>
<td>&gt;1</td>
<td>&lt;1</td>
<td>11</td>
<td>3.6</td>
<td>0.06</td>
<td>0.04</td>
<td>1</td>
</tr>
<tr>
<td>Illumination intensity (W cm⁻¹)</td>
<td>200–400</td>
<td>&lt; 2 × 10⁸</td>
<td>No data</td>
<td>5–10</td>
<td>5</td>
<td>10⁻⁷</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>Depth penetration (µm)</td>
<td>100</td>
<td>10–20</td>
<td>0.1</td>
<td>10–20</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>50⁻⁸</td>
</tr>
<tr>
<td>Multicolour labelling</td>
<td>Unlimited</td>
<td>Limited</td>
<td>Limited</td>
<td>Unlimited</td>
<td>Unlimited</td>
<td>Limited</td>
<td>Unlimited</td>
<td>Unlimited</td>
</tr>
</tbody>
</table>

*Can exceed 150 µm if combined with multiphoton excitation.

Fig. 4. Joint Richardson–Lucy deconvolution of simulated ISM data. (A) 500 × 500 pixel resolution target. Square shows the magnified region with a cluster of point sources. (B) Magnified 25 × 25 pixel view showing cluster of point sources. (C) & (D) Simulated diffraction-limited image. (E) & (F) ISM using the classical excitation. (G) & (H) ISM with doughnut PSFs only.

Planes can be simultaneously measured from a single scan. Using this method to achieve optical sectioning allows for a better axial resolution than confocal microscopy. Combining RESCH with MLD brings some of the lateral resolution increase of ISM and, in simulations, the combination has been shown to produce a 20% increase in resolution in all three dimensions when compared to confocal microscopy (Roider et al., 2016). Although this is only a modest improvement, and some way off a practical technique, this is still an area of ongoing research.
As well as in instrument design, there is significant room for improvement in the image reconstruction process. Currently, the choice of algorithm results in very different reconstructed images, and choosing the right method for a particular application is a complex – at times, subjective – process. Furthermore, MLD is still limited to 2D-SIM reconstruction, and affected by deeper imaging or more optically dense samples. Extending the deconvolution process to include effective consideration of the 3D nature of the sample has not yet been attempted. As well as ignoring the 3D structure in image restoration, MLD also fails to consider the change in the PSF as a function of depth in thicker specimens. This is a known problem in SIM, where changes in the illumination pattern and aberrations in detection result in image artefacts (Booth et al., 2015). Previously correcting for pattern deformation has been addressed using wavefront sensing and adaptive optics (Débarre et al., 2008). Recent work, focusing on computationally accounting for the depth variance of the PSF (Preza & Conchello, 2004; Shaeffitz & Fletcher, 2007; Kim & Naemura, 2015) has shown great promise and application of similar algorithms to MSIM deconvolution and pattern prediction may help to improve the axial resolution and depth penetration. The speed of the reconstruction process has also been an issue, as, unlike methods like STED, the super-resolution image is not immediately available to the user. Current simulations show that a single joint Richardson–Lucy iteration of an MSIM data set (500 × 500 pixel image with 225 pattern shifts) takes approximately 7 s on a modest PC processor. It is likely that advances in processor performance and parallel processing may well allow for deconvolution at similar speed as image capture (Wang et al., 2015).

Acknowledgements

The authors wish to express their gratitude to the Department of Chemistry DTG funding, the Royal Society University Research Fellowship and the Durham Biophysical Sciences Institute for generous funding and scientific contribution with special thanks to Prof. Andrew Beeby for the stimulating scientific discussions over the years.

References


