Methods for optical microscopy – spatial filtering and beyond.

Bright field ; Dark Field ; Phase contrast ; Polarization ; DIC ; Confocal ; Structured illumination ; Sub-diffraction limited imaging (PALM/STORM, STED).

Overview:

Bright field microscopy (absorption, scattering at large angles (outside the collection angle of the objective)

Dark field microscopy (Fourier amplitude filtering)

Phase contrast microscopy (Fourier phase filtering – conversiuon of weak phase modulations to intensity modulations)

Polarization microscopy (analogy with dark field ; birefringent samples)

DIC microscopy (the mach-zehnder interferometer as a converter from phase to intensity ; shearing).

Confocal microscopy (scanning microscopy, real-space amplitude filter, improved depth response, nipkow disk, slit scanning)

Structured illumination microscopy (digital microscopy, deconvolution of images, grating pattern, analogy with moiré)

PALM (stochastic microscopy)

STED, SSIM

Gaussian beams

Start with the paraxial wave equation with a source term

$$\nabla^2 E - \frac{n^2}{c^2}\frac{\partial^2 E}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2 P}{\partial t^2}$$

and assume:

$$E(r,t) = A(r)\exp\left(ikz - i\omega t\right)$$
$$P(r,t) = p(r)\exp\left(ik'z - i\omega t\right)$$

In the slowly varying envelope approximation, this translates to:

$$2ik\frac{\partial A}{\partial z} + \nabla_\perp^2 A = -\frac{4\pi\omega^2}{c^2}\, p\cdot e^{i\Delta kz}$$

One free-space solution to this equation (in 3D) can be written as:

$$A(r,z) = \frac{a}{1+i\zeta}\exp\left(-r^2/w_0^2(1+i\zeta)\right)$$

where ζ is the dimensionless axial coordinate $\zeta = 2z/b$ and $b = kw_0^2$

This can also be rewritten as:

$$A(r,z) = a\frac{w_0}{w(z)}\exp\left(-r^2/w(z)^2\right)\exp\left(ikr^2/2R(z)\right)\exp\left(i\phi(z)\right)$$

with:

$$w(z) = w_0\sqrt{1+\left(\frac{\lambda z}{\pi w_0^2}\right)^2}$$
$$R(z) = z\left(1+\left(\frac{\lambda z}{\pi w_0^2}\right)^2\right)$$
$$\phi(z) = \arctan\left(\frac{\lambda z}{\pi w_0^2}\right)$$

The first exponential term reflects the spatial intensity variation:

$$\left|A(r,z)\right|^2 = a\frac{w_0^2}{w(z)^2}\exp\left(-r^2/w(z)^2\right)^2$$

Clearly, scattering from any plane z is proportional to:

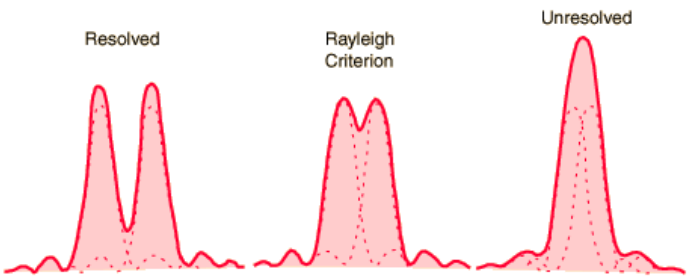
$$\int r dr \left|A(r,z)\right|^2 = \frac{1}{2}\int d\left(r^2\right)\frac{w_0^2}{w(z)^2}\exp\left(-2r^2/w(z)^2\right) = const$$

This merely reflects the conservation of energy through each z plane.

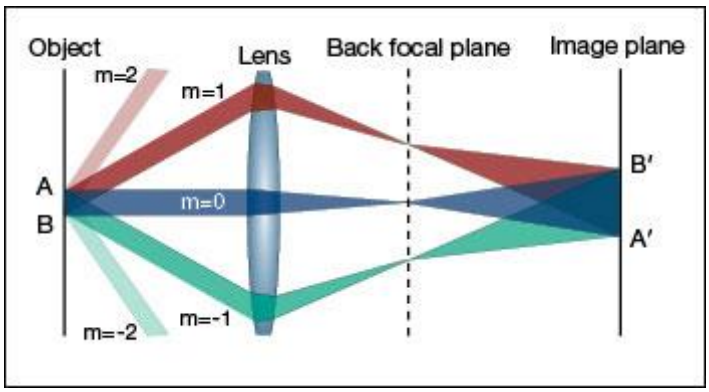
The second term accounts for the field curvature, which is converging at z<0, diverging at z>0 and flat at z=0. The last term is the Guoy phase shift – a π phase shift occurring upon passage through the focus, which is the wave optics analog of the "inside-out" flip in the geometrical focus of a lens.

Optical resolution of a microscope:

The Rayleigh criterion relates to the ability to differentiate two point sources by an imaging lens – and is thus related to the incoherent imaging case. It is relevant, for example, for fluorescent microscopy, where the phase of the emitted light is random. This states that the airy disk pattern of both point sources do not overlap.



The Abbe criterion relates to scattering, or more specifically to the ability to differentiate a grating. It is thus related both to the illumination conditions and the detection conditions. Thus, in most microscopy techniques it is also important to illuminate the sample at a large angle, and a factor of up to 2 in the optical resolution is lost by planar illumination ("coherent vs. incoherent"). This is important in considering how to devise some of the more advanced microscopy techniques used.



Bright field:

This is the simplest imaging technique. It is very commonly used with stained samples, and is intended primarily to detect absorbing species or strongly scattering species.

From a Fourier optics point of view this is effectively imaging only amplitude modulations, where phase modulation (scattering) is only accounted for is scattered light exits the collection angle of the objective. Alternatively, phase objects are revealed upon defocusing. In fact, this is also used to detect whether a scatterer is of higher or lower refractive index than the surrounding.

Dark field.

This is most easily considered in the ASPW approach. The excitation is limited to a narrow ASPW. The simplest realization is with illumination in a plane (DC) component kx=ky=0, and having a stop for the DC component in the Fourier plane of the collection of light through the objective. This enables to image also small angle scattering (unlike bright field). However, using just the DC component results in poorer spatial resolution (considering the effect of illumination angle in the Abbe criterion). Dark field is thus usually realized with a high NA condenser, illuminating through a ring aperture, and a lower NA objective.

Phase contrast

This is a more clever way of converting scattering to intensity modulation. If we set a pi phase shift between scattered light and direct light, weak phase modulations will exhibit themselves as interference-enhanced intensity modulations. Generally resonantly scattered light is at a pi/2 phase shift from the excitation light (just considering a mechanical model of resonant scattering). As such, the intensity at the image plane would be:

$$I_{tot} = \left|A_{direct} + iA_{scattered}\right|^2 = I_{direct} + I_{scattered}$$

The scattered intensity is, however, much smaller than the directly transmitted intensity, which is the reason for the poor contrast of bright filed images. The solution is to shift the relative phase between the two by an extra phase of pi/2, yielding:

$$I_{tot} = \left|A_{direct} \pm A_{scattered}\right|^2 = I_{direct} \pm 2\sqrt{I_{direct} I_{scattered}} + I_{scattered} \approx I_{direct} \pm 2\sqrt{I_{direct} I_{scattered}}$$

This is a specific realization of heterodyne detection, a method for increasing detectibility of a weak coherent signal by interfering it with a strong local oscillator. The relative strength of the local oscillator should be optimized to take into accout two parameters – the dynamic range of the detection system and the noise induced by the local oscillator itself. Since a dynamic range of 3 orders of magnitude is easily

achieved in imaging detectors such as cameras, a 2 order of magnitude signal amplification can be readily achieved.

The outline of a phase contrast system is basically a modified 4f correlator:

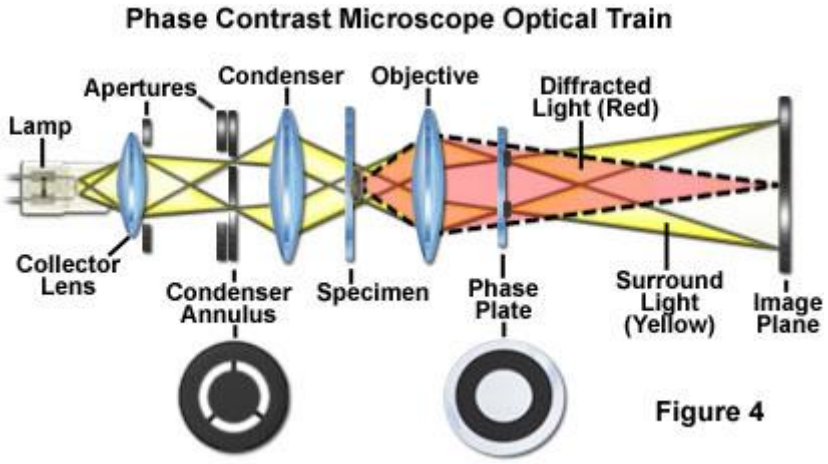


Figure 4

Again, as in dark field, the larger illumination angle contributes to the spatial resolution, hence the extra complication of using an annular aperture rather than a simple circular iris and phase dot at the center of the objective back focal plane.

The details of the scheme are more complicated. The phase plate is also accompanied by an attenuation using a metallic thin film – in order to optimize the contrast and the dynamic range as discussed above. The phase plate can be positive or negative (corresponding to signal appearing as brighter or darker images). Overall, phase contrast is an art by itself, providing much richer data (in general) than brightfield images.

Polarization microscopy

A polarization microscope is a modified bright-field microscope, where the sample is excited by polarized light and imaged through a polarizer. For the case of two parallel polarizers it is analogous to bright field. For two orthogonal polarizers at the input and the output this is analogous to dark field.

Polarization microscopy is mostly used in observing inorganic materials, although some organic materials do exhibit birefringence (most notably collagen).

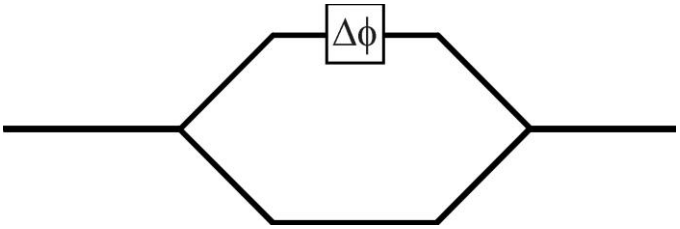
When observing thin specimen the birefringent parts look bright (or dark), in contrast with the background. When the samples are thick, they act like multiple-order waveplates, generating color patterns, depending on which colors are eliminatd from the optical spectrum.

Variations on polarization microscopy include the use of circular polarizers (combinations of polarizer + 1/4 waveplate) to eliminate the dependence on sample orientation), and the addition of birefringent retardation plates to detect the color order (and thus the thickness).

DIC microscopy

DIC is a close cousin of polarization microscopy, where small changes in phase are converted to intensity by use of birefringent elements. However, DIC is a shearing type microscopy, measuring a component of the phase gradient in the sample. The main idea is based on the concept of a mach-zehnder interferometer, with the two arms being neighboring points in the sample.

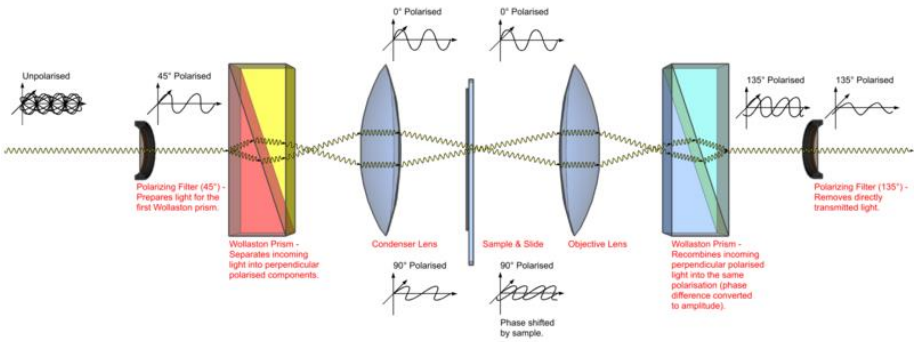
A reminder on a mach-zehnder: Light is split into two arms of equal length. Introduction of a phase shift to one arm results in intensity modulation at the output.



DIC is effectively a shearing mach-zehnder interferometer. The input excitation beam is split into two coherent beams which are effectively focused at neighboring points in the focal plane. Since the two beams overlap spatially in all regions of space other than the focal spot, the only acquired phase difference between them is due to the difference in the optical path in the image plane. After recombination, this yields a shearing image, where intensity is proportional to the phase gradient in the focal plane. Since this is essentially a white light technique, different phase retardations show up as different colors in the image. Often a bias retardation is added to convert the intensity image (observed with no retardation) to a color image.

Blue->orange->red->purple at 200->400->500->600 nm retardation

How is DIC performed in practice:



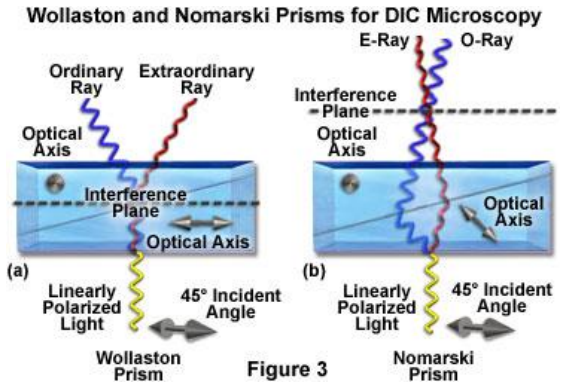


Figure 3

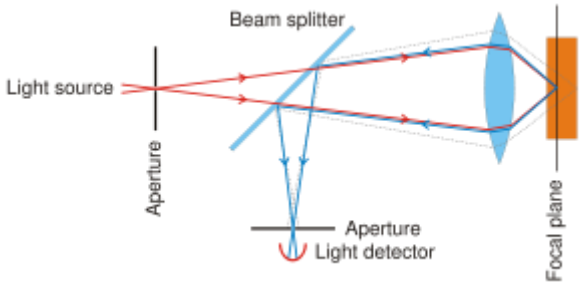
Prisms in DIC

Another advantage of DIC is that it is in practice a sectioning technique, since phase variations arise only from the focal plane. The other common technique for optical sectioning is the confocal microscope.

Confocal microscopy

Out-of-focus objects typically appear as blurred background (effectively noise) in microscopes. The confocal microscope is the most commonly used means to obtain three dimensional images, but does so at a dear price the image is not obtained simultaneously at the entire observation plane but rather through scanning.

The idea, suggested by Minski in 1957, is based on spatial filtering in the real-space domain. A region in 3D space is selected by imaging a point illuminated in the sample onto a pinhole. To make this effective, this is now usually coupled with fluorescence measurements under laser excitation, where various schemes exist to scan the laser rather than the sample (faster). The addition of a confocal pinhole also enables to improve the spatial resolution in the transverse direction.



This can be mathematically described by the following: Consider a 2D sample, and a Gaussian (or Sinc) excitation beam:

$$h(r,z) = h_{exc}(r,z) = a\frac{w_0^2}{w(z)^2}\exp\left(-2r^2/w(z)^2\right)$$

In a standard microscope the detection PSF is unity – any photon can reach the point detector, therefore:

1. The spatial resolution is given entirely by the excitation pattern

2. Out of focus light (in a 3D sample) can also reach the detector. The determining factor there is the shape of the Gaussian beam, whose intensity decreases like the area (z^2), and whose integrated intensity is z-independent.

In a confocal microscope we also add a detection PSF with (generally) a similar shape

$$h_{\det}(r,z)=\frac{w_0^2}{w(z)^2}\exp\left(-2r^2/w(z)^2\right)$$

The overall PSF is the multiplication of the two:

$$h_{tot}(r)=h_{exc}(r)h_{\det}(r)=\frac{w_0^4}{w(z)^4}\exp\left(-\frac{4r^2}{w(z)^2}\right)=\exp\left(-\frac{r^2}{(\sigma/2)^2}\right)$$

Which gives both a rejection in the axial dimension (whose contribution for large z scales now as z^{-4}, and the integrated contribution as z^{-2}), and an increased resolution in the transverse direction.

Here also the use of an engineered excitation light (or confocal pinhole), having a ring shape, can enhance the spatial resolution by the maximal (2X) factor allowed from the Fourier limit (since multiplication in real space = convolution in k-space).

In order to improve the rates of acquisition a variety of spatial multiplexing technique, most notably the Nipkow disk are used. These significantly accelerate image acquisition, alleviating the problems associated with slow scanning.

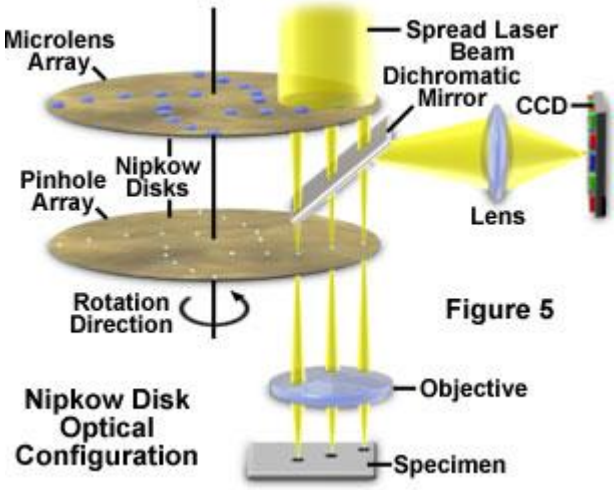


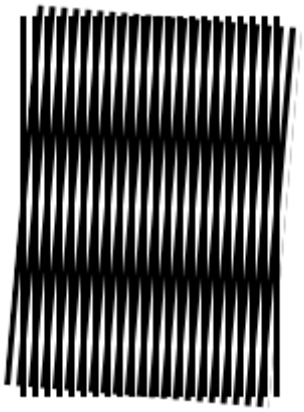
Figure 5

Nipkow Disk Optical Configuration

Structured illumination microscopy

Structured illumination is a Fourier-optics based technique which enables to surpass the Abbe limit by a factor of two at the cost of making the image a completely digital one, utilizing retrieval algorithms, and of an increased complexity.

In the simplest description, structured illumination can be understood using the moiré interference principle. By illuminating the sample with a grating at the spatial resolution limit of the microscope, the moiré fringes from unresolvable gratings up to twice the spatial bandwidth can be observed. The observed grating can be reconstructed using several projections of the excitation grating using a retrieval algorithm. Moreover, this, like DIC, is an optically sectioned widefield technique The main difficulty is that the sample has to be stationary during this set of projection measurements.



In k-space the illumination with a grating basically offsets the transmitted components through the objective by k of the grating (but in it the –k and +k as well as the DC component are mixed, so that several measurements are required to distinguish which observed component belongs where).

Mathematically, for an illumination pattern of the form:

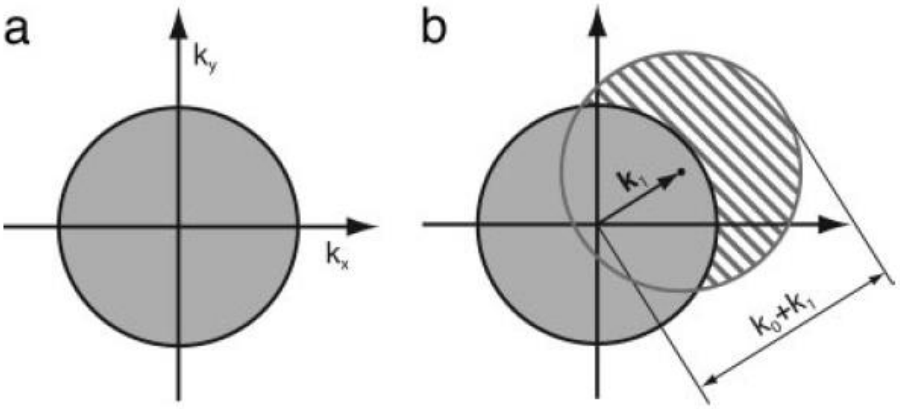
$$I_n(x,y) = I_0\left[1+\cos\left(2\pi\frac{x}{p_x}+2\pi\frac{n}{3}\right)\right]$$

with n=1,2,3. In Fourier space, this leads to:

$$\tilde{I}_n(f_x,f_y) = I_0\left[\delta_x+\frac{1}{2}\left(\delta_x^+\exp\left(-2\pi i\frac{n}{3}\right)+\delta_x^-\exp\left(2\pi i\frac{n}{3}\right)\right)\right]\delta_y$$

with $\delta_x=\delta(f_x)$, $\delta_y=\delta(f_y)$, $\delta_x^+=\delta(f_x+1/p_x)$, $\delta_x^-=\delta(f_x-1/p_x)$

Clearly, the three components $\tilde{I}^{c}, \tilde{I}^{+}, \tilde{I}^{-}$ can be extracted by a linear combination of the three measurements, and by properly choosing the grating period it is possible to synthetically extend the Fourier aperture by a factor of 2 this way.



The main problem of all this: the sample has to be interferometrically (on the scale of λ) stationary during the entire course of the measurement …(hard)

Lensless Imaging

In some cases, the use of an imaging lens is unwanted or simply impossible (VUV or X-ray being the simplest examples of this). In lensless imaging, the Fourier transform of an image is measured, and the inverse transform is performed numerically.

1D phase retrieval problem has no unique solution (analogy with pulse characterization).
2D phase retrieval problem has a unique solution but the retrieval problem is numerically unstable (analogy with FROG).

There are various solutions for this (ptychography, vactorial retrieval). These schemes often use some prior knowledge about the object (‘compact support’) and use interference between different points on the same object or different objects to extract the phases. In many senses, they are reminiscent of holography.

Sub-diffraction limited microscopy

PALM

2D localization based SNR of single emitters is better by sqrt(SNR) than the diffraction limit, enabling in principle unlimited resolution provided:

1. Enough time
2. Reversibly photoactivable fluorophores

3D achieved by introducing an astigmatic element in the detection path so that the Z dimension is encoded in the ellipticity (radius is symmetric for +z and –z).

STED (resolft – Hell)

This is actually a nonlinear microscopy technique (more exactly, a family of nonlinear microscopy techniques), since it relies on saturation of a linear absorption (or stimulated emission) process. However, due to its relative importance in the last several years as the only practical (and commercial) far-field microscope with a sub-diffraction limited PSF for ensembles, it is worth mentioning here.

The basis of the potentially infinite spatial resolution is in the nonlinearity of saturation. If we solve the steady-state for a fluorophores with absorption cross section σ and radiative lifetime τ we get:

$$\frac{dn}{dt} = (1-n)\sigma I - \frac{n}{\tau} = 0 \quad \Rightarrow \quad n = \frac{\sigma I}{\sigma I + \tau^{-1}}$$

Expanding this in a power series for weak excitation we immediately find that an infinite series is required, so that the response is essentially nonlinear with infinite order, and that around saturation the higher order elements dominate.

$$n \propto -\sum_m (-1)^m (\sigma\tau I)^m$$

In practice, the order of nonlinearity for which the coefficients are non-negligible is dependent on the intensity. This is the basis for the enhanced resolution in saturation-based microscopy

In STED and its analogues, a diffraction limited volume is excited, followed by elimination of the excitation by some physical mechanism with a ring-shaped beam, having a single intensity zero at the center of the excitation beam. When this depletion beam is saturated, a smaller and smaller volume is left excited at the center of the original excitation spot.

In the simplest description, the PSF of the STED microscope is given by a multiplication of the excitation profile by the probability to not be depleted:

$$h(r) = h_{exc}(r)\eta(r)$$
$$\eta(r) = \exp(-\sigma\tau\, I_{STED}(r))$$

Taking:

$$h_{exc}(r) = \cos^2\left(\frac{\pi r NA}{\lambda_{exc}}\right)$$

$$I_{STED}(r) = \zeta I_{sat} \sin^2\left(\frac{\pi r NA}{\lambda_{STED}}\right)$$

We get (assuming both wavelengths are comparable):

$$\Delta r \approx \frac{\lambda}{2NA\sqrt{1+\zeta}}$$

Which is a reformulated Abbe criterion, where the resolution is enhanced by the square root of the saturation ratio

SSIM and SAX

In SSIM, the illumination pattern is still:

$$I_n(x,y) = I_0\left[1+\cos\left(2\pi\frac{x}{p_x}+2\pi\frac{n}{3}\right)\right]$$

However, the emission is partially saturated, which gives a nonlinear response as a function of the excitation intensity. Spatially, this generates thin dark stripes in the excitation pattern. By taking more linear combinations, higher order coefficients in k-space can be extracted (provided there is enough amplitude in them – that is that saturation levels are high enough). This is in principle unlimited since for strong saturation the dark stripe thickness becomes a delta function, containing equal contributions of all harmonic orders in K-space.

SOFI

This method utilizes correlations of the emitted radiation. It was invented based on intensity fluctuations of semiconductor nanocrystals used as markers.