Super-resolution fluorescence microscopy

Bernd Rieger

Delft University of Technology

"for the development of super-resolved fluorescence microscopy"

Fluorescence Epi-Illumination

 $\tilde{\mathbf{f}}$ UDelft

Confocal microscopy of thick 3D-samples

Acquisition of Optical Sections

• Scanning microscope with small pinhole on detection side • Removes hazy background from out-of-focus layers

Widefield & confocal fluorescence microscopy

Widefield:

- has uniform illumination
- uses an image sensor for parallel imaging
- is fast
- has poor optical sectioning

Confocal:

- has single spot illumination
- uses a scanner for sequential imaging
- is slow
- has good optical sectioning

Multi-spot scanning microscopy

• Array of spots for video-rate imaging with good optical sectioning

Yokogawa Spinning Disk Unit Optical Configuration

- Drawbacks:
	- light efficiency
	- complex equipment

The diffraction limit.

Optical image of a point object with a lens

- Rays converging to image point/focus correspond to plane waves propagating at angles $\theta<\alpha$
- Total field is the sum of all these plane waves
- Total intensity depends on constructive/destructive interference

Interference gives a spot

- all waves in phase at the image point -> constructive interference/max. intensity
- away from the image point:

waves no longer in phase

- -> destructive interference builds up
- -> intensity decreases

What is the width of the spot?

• Total intensity near zero if waves at largest angles $\pm \alpha$ have destructive interference:

Airy-distribution

• Exact expression for intensity (Point Spread Function) :

$$
I(r) = I_0 \left[\frac{2J_1 (2\pi N Ar/\lambda)}{2\pi N Ar/\lambda} \right]^2
$$

• with $J_1(x)$ a so-called Bessel-function

and the Numerical Aperture is defined by:

 $NA = \sin \alpha$

- Width
- = distance peak to first dark ring:

$$
\Delta r = 0.61 \frac{\lambda}{NA} \frac{1}{f}
$$

Rayleigh

Point Spread Function and Optical Transfer Function

OTF is the Fourier Transform of the PSF

Orders of magnitude

Light microscopy

 $\tilde{\mathbf{f}}$ UDelft

… to optical nanoscopy!

 $\widetilde{\mathbf{T}}$ UDelft

Zoom in & enhance: Unravel the subcelluar machineary of life

 $\tilde{\mathbf{f}}$ UDelft

Super-resolution techniques

- Near-field, surface enhanced
- Far-field

Far-field super-resolution techniques

- Structured Illumintation:
	- Resolution improvement \sim 2x
	- Optical sectioning, 3D possibility
	- Live cell
- Fluorescence switching
	- Resolution improvement \sim 10x
	- Sometimes optical sectioning, 3D possibility
	- Live cell not straight forward

Illumination patterns for resolution increase

How should the illumination change to make only one emitter visible?

Proof of enhanced resolution of SIM images

Widefield SIM

Moiré-effect

- Overlay of periodic patterns gives image with larger period
- Can be used to "magnify" small structures
- = "Structured Illumination Microscopy"

Examples Moiré-effect I

Practical's on Structured Illumination

- Offers 2x increased resolution with good signal transfer of all spatial frequencies.
- To make it work, there are many practical problems:
	- projected grid positions must be known exactly on the sample
	- several grid positions must be acquired
	- the digital reconstruction adds "structured noise"
- "Blind" structured illumination
	- uses many (~ 100) unknown patterns generated by e.g. speckles for illumination.
	- reconstruction must compute sample & illumination

Confocal 2.0

- C.J.R. Sheppard. *Super-resolution in confocal microscopy*. Optik, **1988**
- C.B. Müller and J. Enderlein. *Image scanning microscopy*. Physical Review Letters, **2010**.
- De Luca et al.. *Re-scan confocal microscopy: scanning twice for better resolution*. Biomedical Optical Express, **2013**.
- S. Roth et al. *Optical photon reassignment microscopy (OPRA)*. Optical nanoscopy, **2013**.
- York et al. *Instant super-resolution imaging in live cells and embryos via analog image processing*. Nature Methods, **2013**.

Confocal 2.0

Confocal 2.0

- Offers increased resolution
- Theoretical 2x better, but limited by strong decay of the OTF practically only 1.4x better.
- Very easy to retro-fit on existing confocal microscopes.

STimulated Emission Depletion (STED)

Stimulated emission

Incident photon "knocks out" molecule in excited state Emitted photon identical to incident photon (energy, direction of propagation, polarization)

Application of stimulated emission in STED

Excite fluorophores with first spot Illuminate with second ring-shaped spot Deplete excited state via stimulated emission Collect fluorescent light from central "spike"

STED-setup EXC spot **STED spot** Lens $+$ Detector Effective fluor. spot (PSF) **EXC STED** beam beam 66_{nm} phase Piezo modulation Stage -200 0.200 -200 0 200 $[nm]$

- STED-beam must have bit larger wavelength (Stokes-shift)
- STED-spot must be engineered to "doughnut"-beam/ring-shaped spot
- Needs high powers to fully deplete excited state

source: http://www.mpg.de/english/illustrationsDocumentation/documentation/pressReleases/2006/pressRelease20060412/index.html

 0.5

What are typical powers needed?

Probability that photon de-excites fluorophore by stimulated emission must be very close to one

#photon-hits = intensity \times cross-section \times excited state lifetime

Resolution of STED

Increase in depletion beam power => narrower emission spike => better resolution

Modified Abbe-formula:

$$
d = \frac{\lambda}{2n \sin \alpha \sqrt{1 + I/I_{STED}}}
$$

V. Westphal, S.W. Hell (2005), *Phys. Rev. Lett.* **94**, 143903.

.

STED

Idea:

- "Switch-off" the emitters in the PSF expect at the very center.
- Switching mechanism: Prevent the emitters from emitting.

Idea is around since 1990s, took a long time to make it work See literature list for more

Grotjohan et al., Nature 2011

Figure 2 | Rewritable data storage. The text of 25 Grimm's fairy stories (ASCII code; 1.9 Mbits) consecutively written and read on a 17×17 µm area of a PAA layer containing rsEGFP, with bits written as spots (representative frames shown). The white dots mark spots that were recognized as set bits ('1's). The graph shows an intensity profile along the indicated area, averaged over three pixels along the y-axis. The blue line indicates the threshold used to assign read spots to '0's or '1's.

"PALM" and "STORM"

 $\tilde{\mathbf{f}}$ UDelft

Single emitter localization

Switch fluorophores "on" or "off" **in time**

Good ides iff: Position of single emitter can be determined with high certainty

Mechanisms for "on" / "off" switching

PALM = Photo-Activation Localization Microscopy

Betzig/ H.Hess, Science 2006

S. Hess, Biophysical Journal 2006

PALM workflow

STORM = STochastical Optical Reconstruction Microscopy (Zhuang, Nat. Methods 2006)

Mechanisms for "on" / "off" switching

GSDIM = Ground State Depletion followed by single molecule IMaging (Hell, PRL 2007)

Mechanisms for "on" / "off" switching

dSTORM = "direct" STORM (Heilemann, Angew. Chem. 2008)

Why is a single emitter better than many?

• Single emitters' positions are estimated with nanometer precision

A real acquisition on a microscope

Pointillism in art

Dithering for paper printing

http://en.wikipedia.org/wiki/Dither

World press photo 2012

Nanoscopy with more than 100,000 'doughnuts'

RESOLFT =REversible Saturable Optical Fluoresece **Transitions**

= nonlinear structured illumintation

 $I_{\text{OH}} = I \sin^2(x) + I \sin^2(y)$

C

 $III_{-} = 5$

 $III_n = 100$

Super-resolution ~10-20 nm

K.Xu, G. Zhong, X. Zhuang, Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons, **Science** 339: 452, 2013.

Segmentation

Series of Uniform Filter operation

Local maximum

 $\widetilde{\mathbf{f}}_{\text{UD}}$

Localization uncertainty: rule-of-thumb

• Position estimate by each photon with standard deviation:

$$
\sigma \approx \frac{\lambda}{NA}
$$

- Independent estimates for N_{ph} photons
- The overall standard deviation is then:

$$
\Delta x \approx \frac{\lambda}{NA\sqrt{N_{ph}}}
$$
\n
$$
\text{typical number } (\lambda = 488 \text{ nm},
$$
\n
$$
\text{NA} = 1.25, \text{N}_{\text{photons}} = 400)
$$
\n
$$
\Delta x = 20 \text{ nm}
$$

Imaging Model

2D finite pixel Gaussian PSF

$$
PSF(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x-\theta_x)^2 - (y-\theta_y)^2}{2\sigma^2}}
$$

$$
\mu_k(x, y) = \theta_{I_0} \int_{A_k} PSF(u, v) du dv + \theta_{bg}
$$

$$
\mu_k(x, y) = \theta_{I_0} \Delta E_x(x, y) \Delta E_y(x, y) + \theta_{bg}
$$

$$
\Delta E_x(x, y) \equiv \frac{1}{2} \text{erf}\left(\frac{x - \theta_x + \frac{1}{2}}{2\sigma^2}\right) - \frac{1}{2} \text{erf}\left(\frac{x - \theta_x - \frac{1}{2}}{2\sigma^2}\right)
$$

$$
\Delta E_y(x, y) \equiv \frac{1}{2} \text{erf}\left(\frac{y - \theta_y + \frac{1}{2}}{2\sigma^2}\right) - \frac{1}{2} \text{erf}\left(\frac{y - \theta_y - \frac{1}{2}}{2\sigma^2}\right)
$$

Ignore: vectorial effect, aberrations dipole orientation, readout noise

Gaussian is OK for free rotating emitter

Parameters:

q*x*: x-position q*y*: y-position q*I0*: emission rate q*bg*: background count rate q*s*: spot width

µ_*k*: expected count in pixel *k*

Probability and likelihood

Probability of observing pixel values $\{n_k|k = 1,...,N_{pix}\}$ given average rate $\{\mu_k|k = 1,...,N_{pix}\}\$ is given by Poisson law:

$$
P(\lbrace n_k \rbrace | \lbrace \mu_k \rbrace) = P(n_1 | \mu_1) P(n_2 | \mu_2) ... P(n_{N_{pix}} | \mu_{N_{pix}})
$$

$$
P(n_j | \mu_j) = \frac{e^{-\mu_j} \mu_j^{n_j}}{n_j!}
$$

The reverse is also true: *Likelihood* that $\{\mu_k | k = 1,...,N_{pix}\}$ corresponds to ground truth given the observed pixel values $\{n_k|k = 1,...,N_{pix}\}$ is:

$$
L\left(\{\mu_k\}\Big|\{n_k\}\right) = L\left(\mu_1\Big|n_1\right)L\left(\mu_2\Big|n_2\right)...L\left(\mu_{N_{pix}}\Big|n_{N_{pix}}\right) \nL\left(\mu_j\Big|n_j\right) = \frac{e^{-\mu_j}\mu_j^{n_j}}{n_j!}
$$

MLE optimization problem statement

The unknown parameters are found for the maximum in the likelihood L({ μ_k }|{ n_k }) as a function of the unknown parameters (q_1, q_2, q_3, q_4) = (x_0, y_0, σ, N) .

Maximum of $L({\mu_k}|\{n_k\}) =$ maximum of log-likelihood:

$$
\log \left(L\left(\{\mu_k\} \big|\{n_k\}\right)\right) = \sum_{k=1}^{N_{pix}} \log \left(L\left(\mu_k \big| n_k\right)\right) \\ = \sum_{k=1}^{N_{pix}} \left(n_k \log \left(\mu_k\right) - \mu_k - \log \left(n_k\right)\right)
$$

Then we must solve:

$$
0 = \frac{\partial \log \left(L\left(\{\mu_k\}|\{n_k\}\right)\right)}{\partial \theta_j} = \sum_{k=1}^{N_{pix}} \left(\frac{n_k}{\mu_k} - 1\right) \frac{\partial \mu_k}{\partial \theta_j} \quad \text{for } j = 1, 2, 3, 4
$$

Finding the optimum

Using that:

$$
\frac{\partial \log(L)}{\partial \theta_j} = \sum_{k=1}^{N_{pix}} n_k \frac{\partial \log(\mu_k)}{\partial \theta_j} - \frac{\partial \left(\sum_{k=1}^{N_{pix}} \mu_k\right)}{\partial \theta_j} \approx \sum_{k=1}^{N_{pix}} n_k \frac{\partial \log(\mu_k)}{\partial \theta_j} - \frac{\partial N}{\partial \theta_j}
$$

$$
\log(\mu_k) = -\frac{(x_k - x_0)^2 + (y_k - y_0)^2}{2\sigma^2} + \log(N) - 2\log(\sigma) - \log(2\pi)
$$

we find:

$$
N = \sum_{k=1}^{N_{pix}} n_k, \quad x_0 = \frac{1}{N} \sum_{k=1}^{N_{pix}} n_k x_k, \quad y_0 = \frac{1}{N} \sum_{k=1}^{N_{pix}} n_k y_k
$$

$$
\sigma^2 = \frac{1}{2N} \sum_{k=1}^{N_{pix}} n_k \left(\left(x_k - x_0 \right)^2 + \left(y_k - y_0 \right)^2 \right)
$$

2nd derivative at optimum is measure for uncertainty:

$$
\frac{\partial^2 \log(L)}{\partial x_0^2} = \sum_{k=1}^{N_{pix}} n_k \frac{\partial^2 \log(\mu_k)}{\partial x_0^2} = -\frac{N}{\sigma^2}
$$

This gives an uncertainty in emitter location:

$$
\Delta x_0 = \left[-\frac{\partial^2 \log(L)}{\partial x_0^2} \right]^{-1/2} = \frac{\sigma}{\sqrt{N}}
$$

Likelihood and CRlB

 $L(\vec{x}|\theta) = \prod_{k} \frac{\mu_k(x,y)^{x_k} e^{-\mu_k(x,y)}}{x_k!}$ Poisson process

Cramer-Rao lower bound & Fisher information matrix

$$
var(\hat{\theta}) \ge I(\theta)^{-1} \qquad I_{ij}(\theta) = E\left[\frac{\partial \ln(L(\vec{x}|\theta))}{\partial \theta_i} \frac{\partial \ln(L(\vec{x}|\theta))}{\partial \theta_j}\right]
$$

$$
I_{ij}(\theta) = \sum_k \frac{1}{\mu_k(x, y)} \frac{\partial \mu_k(x, y)}{\partial \theta_i} \frac{\partial \mu_k(x, y)}{\partial \theta_j}
$$

Use CRIb to estimate localization uncertainty

(CRLB: Minimum variance of an estimated parameter)

66

Parallel processing on Graphics Card

fixed number of iterations fits GPU's single instruction multiple data strategy (SIMD)

>100 processors on one GPU, price 100-1000€

Size of sub-regions fixed on $2\sigma_{\text{pSF}} + 1$ pixel

Performance on Simulations

RE Thompson, DR Larson & WW Webb, Biophysical Journal 82, 2002

Post-processing of localizations (1)

• Filter the localizations based on the **fitted parameters** e.g. photon count, precision, width of the PSF

Post-processing of localizations (2)

- "Frame connection": Combining spatially close localization from consecutive frames.
- Higher photon count for the individual emitter leads to higher effective localization precision.
- Danger of merging two different emitters into one effective localization.

Post-processing of localizations (3)

• Drift correction

- Imaging times are typically 5-30 minutes! Drift on the order of the localization precision \sim 10 nm must be avoided!
- 1) Use reference beads for tracking
- 2) Use localizations for cross-correlation based drift correction
- M.J. Mlodzianoski et al. Sample drift correction in 3d fluorescence photoactivation localization microscopy. Optics Express, 19(16): 15009-15019, 2011.
- C. Geisler et al., Drift estimation for single marker switching based imaging schemes. Optics Express, 20(7):7274-7289, 2012.

Resolution in localization microscopy

Resolution criteria in super-resolution

Currently used measures for localization microscopy

σ

- Nyquist sampling density: $2/\sqrt{\rho}$
- Localization precision:

Fourier Ring Correlation (FRC) resolution Qualitative validity in experiments

• Actin filaments, Alexa647 coupled to Phalloidin

Widefield Binned localizations FRC resolution = 100 nm FWHM loc. unc. \sim 38 nm

Fourier Ring Correlation (FRC) resolution

Qualitative validity in experiments

• Actin filaments, Venus4 label

Data courtesy of Kees Jalink & Daniela Leyton Puig

Fourier Ring Correlation

different options for actual splitting of the time series

Qualitative validity in simulations

Qualitative validity in simulations

• The green circles show where the distance between the arms is equal to FIRE

TUDelft

Resolution as a function of acquired frames

Acquisition of tubulin in HeLa cells with Alexa 647

TUDelft

Effect of fitting algorithms

dSTORM acquisition of tubulin in HeLa cells with Alexa 647

- Maximum Likelihood
- Least Squares
- Centroid

Effect of stage drift

GSDIM acquisition of actin filaments; HeLa cells (Alex647 coupled to Phalloidin)

Data courtesy of Daniel Leyton Puig & Kees Jalink

Drift corrected \sim 70 nm

Resolution prediction

• Expected value of the correlation curve:

$$
\langle FRC(q) \rangle = \frac{\sum_{\vec{q} \in \text{circle}} \left(Q + \mathbb{N} |\hat{\mathbf{V}}(\vec{q})|^2 \right) \exp\left(-4\pi^2 \sigma^2 q^2 \right)}{\sum_{\vec{q} \in \text{circle}} \left[2 + \left(Q + \mathbb{N} |\hat{\mathbf{V}}(\vec{q})|^2 \right) \exp\left(-4\pi^2 \sigma^2 q^2 \right) \right]}
$$
\nNumber of
localizations shape

\nAverage localizations per molecule;

\nQ=0, each emitter is only seen once

 $\tilde{\mathbf{T}}$ UDelft

• Expected resolution for 2 sinusoidal lines:

$$
\langle FRCresolution \rangle \approx \frac{2\pi\sigma}{\sqrt{W(6\pi\rho\sigma^2)}}
$$

W(x) Lambert W-function

Density or precision limited?

 $\tilde{\mathbf{T}}$ UDelft

Density or precision limited?

Density or precision limited?

Acquisition of tubulin in HeLa cells

FRC computed as local measure

Anisotropic imaging

• Generalization of FRC to anisotropic data by correlating over lines (2D) or planes (3D) instead of rings or spherical shells

Anisotropic imaging in 3D

Anisotropic imaging in 2D

Fourier Line Correlation

 $\tilde{\mathbf{f}}$ UDelft

Relation to classical resolution

- In localization microscopy, FIRE does not reduce exactly to the Nyquist resolution in the limit of high photon counts
	- For two lines:

$$
\lim_{\sigma \to 0} \langle FIRE \rangle = \sqrt{\frac{2\pi}{3\rho}} \approx 1.45 R_{Nyquist}
$$

• In widefield microscopy, FIRE reduces to Abbe limit for low noise conditions

FRC resolution in widefield microscopy

Widefield acquisition of 200 nm fluorescent beads; just repeated imaging of the same sample

60x 0.7 NA lens 0.49x Nyquist pixel size

Estimated spurious correlation Q via model

• Idea: use the expected FRC to estimate the spurious term

$$
\langle FRC(q) \rangle = \frac{\sum_{\vec{q} \in \text{circle}} \left[Q \exp\left(-4\pi^2 \sigma^2 q^2 \right) + \left[N \hat{\psi}(\vec{q}) \right]^2 \exp\left(-4\pi^2 \sigma^2 q^2 \right) \right]}{\sum_{\vec{q} \in \text{circle}} \left[2 + \left(Q + N \hat{\psi}(\vec{q}) \right]^2 \exp\left(-4\pi^2 \sigma^2 q^2 \right) \right]}
$$
\nDepends on average localizations per detected localizations shape

\nlocalizations per detected and localizations shape on the original data. The total values are given by the total values of the total values of the total values.

Plan:

- Only calculate the numerator
- Divide the exponential away (with estimation)

Estimated spurious correlation Q via model

• Idea: use the expected FRC to estimate the spurious term

$$
\langle FRC(q) \rangle = \frac{\sum_{\vec{q} \in \text{circle}} \left[Q \exp\left(-4\pi^2 \sigma^2 q^2 \right) + N |\hat{\mathcal{W}}(\vec{q})|^2 \exp\left(-4\pi^2 \sigma^2 q^2 \right) \right]
$$

Number of Object localizations

shape

Localization precision

Plan:

- Only calculate the numerator
- Divide the exponential away (with estimation)

Estimated spurious correlation Q via model

• Idea: use the expected FRC to estimate the spurious term

Number of Object localizations shape

Localization precision

Plan:

- Only calculate the numerator
- Divide the exponential away (with estimation)

Data courtesy of Mark Bates

Two color data of tubulin

Data courtesy of Mark Bates

How to display localization images anyway?

- Localization microscopy produces data but no images
- Estimated fit parameters:
	- $x, y, z)$ position
	- localization uncertainty
	- intensity
	- background
	- goodness of fit/ Fisher information
- Reconstruction of the object in the Nyquist sense is not considered part of the visualization process

Histogram binning [Egner et al. Biophysical Journal 2007] Gaussian rendering [Betzig et al. Science 2006] Jittered histogram binning [Krizek et al. Optics Express 2011] Delaunay triangulation & Quad-tree visualization [Baddeley et al. Microscopy and Microanalysis 2010]

Considerations for visualization

• Intuitive interpretation:

Linearity of intensity values with emitter density

[not preserved by scattergram, Triangulation, Quadtree]

• Give the **best possible image resolution**.

 $\tilde{\mathbf{f}}$ UDelft

• Gaussian weights localizations to q depending their σ Imprecise localizations decrease faster as function of q

UDelft

(ignoring pixelation effect on histogram)

Conclusions: visualization

- Gaussian rendering is best, especially if the localization error is large (theoretical proven)
- Gaussian rendering only better than histogram binning if each emitter is rendered with its own localization uncertainty
- Gaussian, jittering and histogram binning give typically same resolution
- Quad-tree and triangulation give irregular bias and should not be used. Especially for low densities.

From Q to emitter counting

$$
Q = \langle M \rangle + \left(\frac{\text{Var}\left(M\right)}{\langle M \rangle} - 1\right)
$$

- Q can be used to estimate $\langle M \rangle$ = the average number of times one molecule is localized
- This requires a model for emitter switching:

Statistical switching model

Expected number of localizations per emitter M vs. Q

 $\tilde{\mathbf{T}}$ UDelft

The unknown rate constant

 k_{eff} is estimated from cumulative localizations

$$
\langle M \rangle = \left(1 + \frac{k_{\text{off}}}{k_{\text{b}}}\right) \left(1 - \exp\left(-\frac{k_{\text{eff}}t}{k}\right)\right) \approx N
$$

Fit model: $y = a(1 - \exp(-bt))$

Results: $a = 4.8 10⁴$ b = 4.0 10^{-4} \Longrightarrow k_{eff} $R^2 = 0.9987$

Control experiment with ground truth

Sparse Alexa 647 labeled antibodies on a glass surface

Control experiment with ground truth

Photobleaching effects are correctly included

 $\tilde{\mathbf{T}}$ UDelft

Application to dual color data

Estimating the unknown rate constant

Application to dual color data

of localizations per emitter and all rate constants

Conclusions

- FRC Image Resolution is proposed as a image based resolution measure for localization microscopy
- Sensitive to labelling density, localization precision, sample shape
- Can deal with multiple localizations per emitter
- Counting of average localizations per emitter
	- \checkmark Without prior knowledge of sample structure
	- \checkmark Without a calibration experiment for rate constants
- Neglecting photobleaching leads to overestimation
	- Also for pair-correlation approach
- Caveats:
	- Transition rates are assumed constant
	- False positive localizations
	- Blinking/bleaching model may be too simplistic

