Optimizing Point Spread Functions to Discern Highly Overlapping Emission Spectra

Sanduni Fernando[1], Jason Martineau[2], Erik Jorgensen[3] and Jordan Gerton[2]

[1]Department of Physics and Astronomy, University of Utah, Salt Lake City, USA, SALT LAKE CITY, Utah, United States, [2]Department of Physics and Astronomy, University of Utah, Salt Lake City, USA, United States, [3]Department of Biology, University of Utah, Salt Lake City, USA, United States

Point Spread Function (PSF) engineering is a branch of Localization Microscopy (LM) that can improve localization precision and allow identification of multiple spectral channels without using traditional emission filters or spectrometers.

We and others [1-4] have shown that incorporating a phase mask at a pupil plane in the fluorescence detection path introduces structure to the Airy PSF, which facilitates spectral discernibility and/or spatial localizability [5]. However, a formalism to obtain a phase mask that yields optimum spectral discernibility for a given set of spectra has not been discussed widely. We demonstrate four figures of merit (FOM) below, derived under different veins of thought to optimize a phase element given a set of emission spectra. This work builds on previous results where we introduced the Pixel Confusion FOM [6].

To develop all our FOM, we consider PSFs limited by Poisson noise. These PSFs occupy the same central spatial coordinates and are represented by a particular number of pixels on the camera. We compare pairs of PSFs corresponding to each spectrum for all possible permutations among spectra at each camera pixel within the region of interest. We take the sum along all pixels to represent the distribution of photons within the whole region. Then we maximize the FOM (other than Pixel Confusion, which is minimized) to maximize the difference between the PSFs. We assume that all fluorophores have nearly equal brightness, i.e., the average photon detection rate for all emission spectra is the same.

The first FOM, Kullback-Leibler Divergence score (KLD), also known as relative entropy [7] measures the divergence of one distribution from another, or from an information theory perspective, states the quantity of extra information required to sufficiently represent one distribution in the presence of another. For two Poisson distributions that need to be distinguished from each other, KLD is expressed as:

$$\text{KLD}_i = (\mu_{1i} - \mu_{2i})\log\left[\frac{\mu_{1i}}{\mu_{2i}}\right]$$

where μ_{1i} and μ_{2i} are the values of the model PSFs for two spectra at the ith pixel. Maximizing KLD within an upper and a lower bound for thicknesses yields an optimum configuration of thicknesses for a phase element within the bounds.

The second FOM, Chernoff Information is defined as the best achievable exponent in Bayesian probability of error. It is a measure of how hard it is to separate two PSF signals into two classes by performing a hypothesis test on them. For two Poisson distributions, Chernoff Information is in a closed-form as given by C_i [8].

$$C_i = \mu_{1i} \frac{\left(\frac{\mu_{2i}}{\mu_{1i}} - 1\right)\left(\log\left(\frac{\frac{\mu_{2i}}{\mu_{1i}} - 1}{\log\frac{\mu_{2i}}{\mu_{1i}}}\right) - 1\right) + \log\frac{\mu_{2i}}{\mu_{1i}}}{\log\frac{\mu_{2i}}{\mu_{1i}}}$$

$$a = \mu_{2i}/\mu_{1i}$$

We maximize C_i to optimize the thicknesses.

The third FOM, the square of the difference between two PSF models (SQD), is a maximum when the two PSFs are very different from each other. This difference can be rearranged in terms of arithmetic and geometric means such that for two Poisson distributions, this is equal to the Bhattacharyya Distance:

$$SQD_i = \frac{\mu_{1i} + \mu_{2i}}{2} - \sqrt{\mu_{1i}\mu_{2i}}$$

Maximizing this distance is a way to obtain phase element thicknesses that enhances spectral discernibility.

The fourth FOM, is a heuristic based on the Skellam distribution, which is the discrete probability distribution of the difference of two statistically independent Poisson-distributed random variables (number of photons in this context) with expected values μ_{1i} and μ_{2i}[9]. It would be skewed when μ_{1i} and μ_{2i} are unequal. For a certain difference in photons k, the probability would be maximum when the distribution is narrow and less tailed. Therefore, combining the mean ($\mu_{1i} - \mu_{2i}$), variance ($\mu_{1i} + \mu_{2i}$), and the kurtosis ($3 + 1/(\mu_{1i} + \mu_{2i})$) of the distribution we obtain a heuristic FOM, given by S_i.

$$S_i = \frac{mean^\alpha}{kurtosis^\beta \times variance^\gamma}$$

We assigned integer values for the factors α, β, γ and carried out the optimization for each combination of these factors. We obtained the best optimization at $\alpha=4, \beta=2, \gamma=3$:

$$S_i = \frac{(\mu_{1i} - \mu_{2i})^4}{(3(\mu_{1i} + \mu_{2i}) + 1)^2 (\mu_{1i} + \mu_{2i})}$$

We maximized $\log(S_i)$ to optimize the phase element thickness configuration.

We used the five FOM (four described above and Pixel Confusion [6]) to generate four-quadrant SiO_2 phase plates with optimized thicknesses for highly overlapping spectra of Alexa Fluor 647, Alexa Fluor 660, and Alexa Fluor 680 with a thickness upper-bound of 6 μm and a lower-bound of zero. The PSFs (shown in Figure 1) were generated using the Gibson-Lani model [10] assuming 1,000 photons for each spectrum (nearly equal brightness). Then, we assessed the performance of these phase plates using Monte-Carlo simulations with differing photon budgets (shown in Figure 2).

All the above FOM have closed forms and are independent of the number of photons which makes them computationally efficient. One FOM may outperform the other and provide better spectral discernibility depending on the overlap between spectra, wavelength range, thickness upper bound for the phase plate,

and/or the ratio between the emitter brightness. It is important to note that these FOM are intended to be used when all the emitters have nearly equal brightness.

For the four-quadrant SiO_2 phase plate, the above spectra with the thickness upper-bound of 6 μm, Monte-Carlo simulations suggest that Chernoff Information and KLD FOM yield higher-performing phase plates compared to Pixel Confusion. SQD and the heuristic based on Skellam Distribution FOM give comparable results to that of Pixel Confusion. We notice that when the thickness upper-bound is 6.5 μm and 10,000 photons are used for optimization, the results converge to similar phase plate thickness configurations for all FOM. These values are the same as that were obtained using Chernoff Information FOM at a thickness upper-bound of 6.5 μm and only 1,000 photons for optimization, which suggests that Chernoff Information is the most robust FOM for optimizing the phase plate configuration.
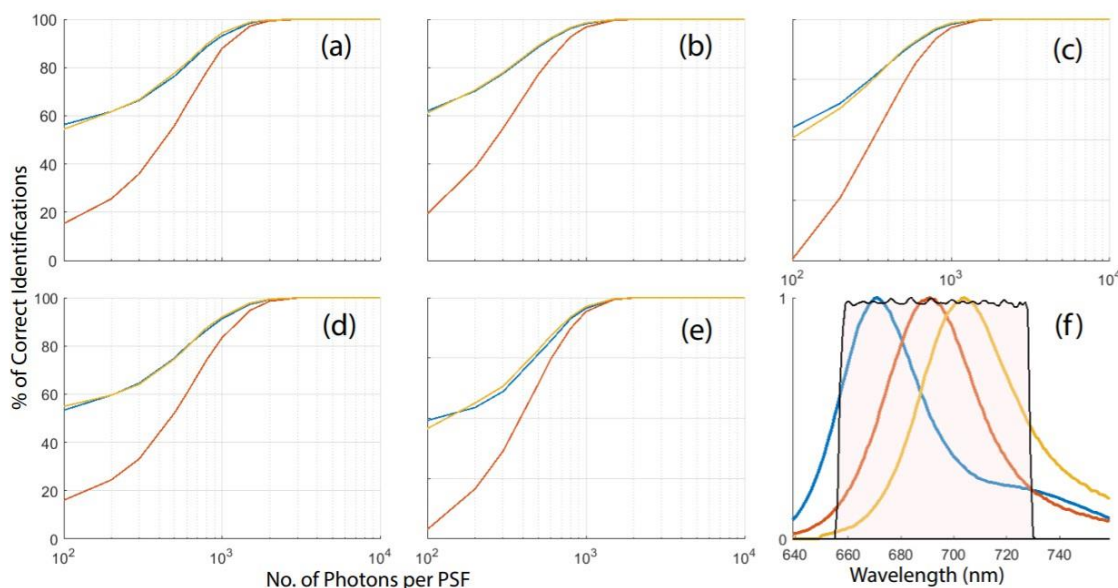

Figure 1. The PSFs from each FOM and their corresponding Phase Plate (PP) thicknesses. Panel (a) Pixel Confusion, (b) Chernoff Information, (c) KLD, (d) SQD, and (e) Skellam Distribution. Each panel consists of the PSFs for Alexa Fluor (AXF) 647, 660, and 680 from top to bottom respectively with a common color bar indicating pixel value range. The corresponding quadrant PP thickness configuration in μm is shown at the bottom of each panel. The first quadrant of the PP is held at zero. The colored circle indicates the beam extent.
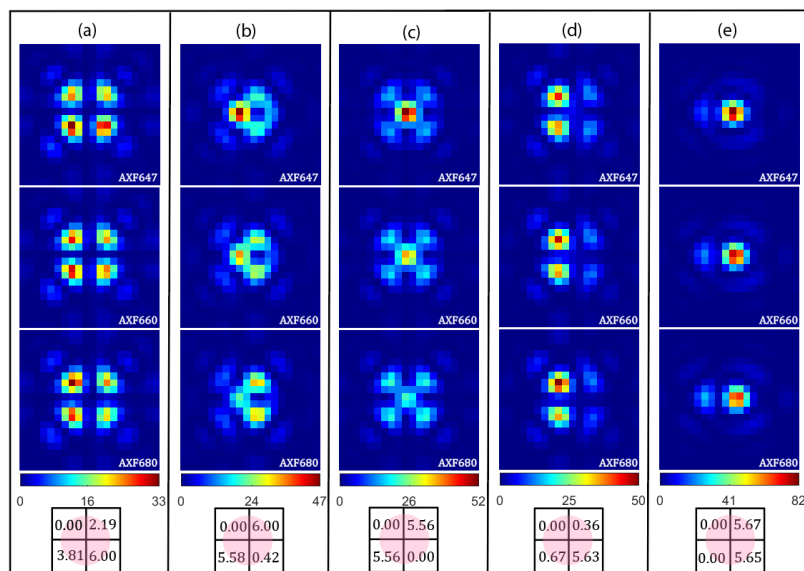
Figure 2. Monte-Carlo simulation results for different photon budgets. Panels (a)-(e) correspond to the Monte-Carlo results for the PP resulting from Pixel Confusion, Chernoff Information, KLD, SQD, and Skellam Distribution FOM respectively. The horizontal axis is the number of photons per PSF and the vertical axis is the correct number of spectra identifications as a percentage of total simulations. Blue, red and yellow lines correspond to Alexa Fluor 647, 660, and 680 respectively. All PSF families show near or over 90% spectral discernibility at 1,000 photons and around 100% above 2,000 photons. Panel (f) shows the normalized spectra of the dyes Alexa Fluor 647(blue), 660(red), and 680(yellow) in the presence of a Semrock Penta Band dichroic spectral filter. The shaded region in the image indicates the region of the spectra used for optimization and Monte-Carlo simulations.

References

[1] Y. Shechtman, L. E. Weiss, A. S. Backer, M. Y. Lee and W. E. Moerner, "Multicolour localization microscopy by point-spread-function engineering," Nature Photonics, vol. 10, no. 2016, p. 590–594, 2016.

[2] Y. Shechtman, L. E. Weiss, A. S. Backer, S. J. Sahl and W. E. Moerner, "Precise Three-Dimensional Scan-Free Multiple-Particle Tracking over Large Axial Ranges with Tetrapod Point Spread Functions," Nano Letters, vol. 15, no. 6, p. 4194–4199, 2015.

[3] J. Broeken, B. Rieger and S. Stallinga, "Simultaneous measurement of position and color of single fluorescent emitters using diffractive optics," Optics Letters, vol. 11, no. 3352-3355, p. 39, 2014.

[4] Y. Zhou, V. Zickus, A. R. Harvey and P. Zammit, "Double-Cubic Point Spread Function for 3D Extended-Depth Localization Microscopy," in Imaging and Applied Optics, 2018.

[5] Y. Shechtman, S. J. Sahl, A. S. Backer and W. Moerner, "Optimal Point Spread Function Design for 3D Imaging," Physical Review Letters, vol. 113, no. 13, p. 133902, 2014.

[6] J. T. Martineau, R. Menon, E. M. Jorgensen and J. Gerton, "A Proposed Method for Optimizing the Spectral Discernibility of Engineered Point-spread Functions for Localization Microscopy," in Microscopy and Microanalysis , Volume 25 , Supplement (S2) , pp. 1232 - 1233, 2019.

[7] J. Brownlee, "How to Calculate the KL Divergence for Machine Learning," 18 October 2019. [Online]. Available: https://machinelearningmastery.com. [Accessed 20 December 2020].

[8] F. Nielsen, "Chernoff information of exponential families," arXiv 1102.2684v1 , 2011.

[9] A. A. Alzaid and M. A. Omair, "On The Poisson Difference Distribution," The Bulletin of the Malaysian Mathematical Society, 2009.

[10] H. Kirshner, F. Aguet, D. Sage and M. Unser, "3-D PSF fitting for fluorescence microscopy: implementation and," Journal of Microscopy, vol. 249, no. Pt 1 2013, p. 13–25, 2012.