

Supplemental document accompanying submission to *Journal of the Optical Society of America A*

Title: RobustandUnbiasedEstimationoftheBackground
DistributionforAutomatedQuantitativeImaging

Authors: Hernán Grecco, Mauro Silberberg

Submitted: 10/5/2022 5:22:44 PM

OPTICA
PUBLISHING GROUP
Formerly OSA

Robust and Unbiased Estimation of the Background Distribution for Automated Quantitative Imaging: supplemental document

1. HIGH-FREQUENCY SAW-LIKE STRUCTURE YIELDS SMALL SMO VALUES

As the derivative changes direction faster than the averaging kernel, it yields a null SMO value and therefore indistinguishable from uncorrelated noise.

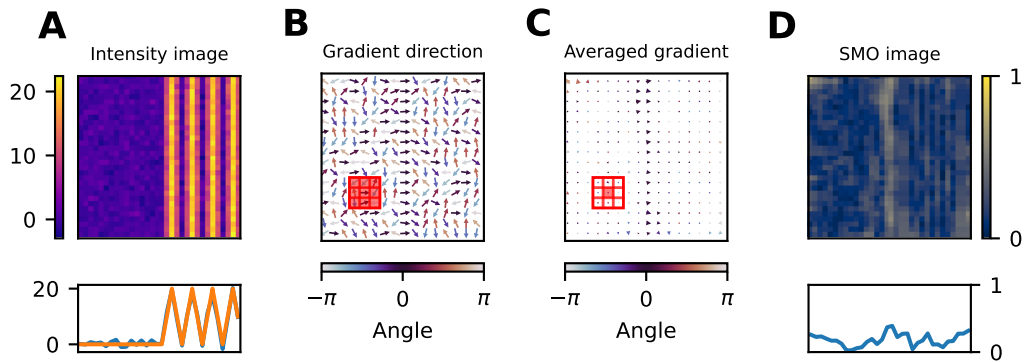


Fig. S1. Breakdown of the Silver Mountain Operator. A) Intensity image simulated as constant intensity followed by a saw-like region, plus standard normal noise. On the bottom row, a line profile along a row before (orange) and after (blue) adding noise. B and C) Gradient direction and its local average, respectively, color-coded by angle. The red square indicates the averaging kernel. D) Resulting SMO image, which is the length of the average gradient.

2. ROBUSTNESS OF FOR DIFFERENT AVERAGING KERNEL SIZES

The choice of averaging kernel size does not have a great impact on the recovered background distribution, as long as it is smaller than the foreground structures, since they would be averaged out. Nevertheless, larger kernels do produce a slightly better estimate, as more of the neighbourhood is considered to decide if a pixel is part of the background or not.

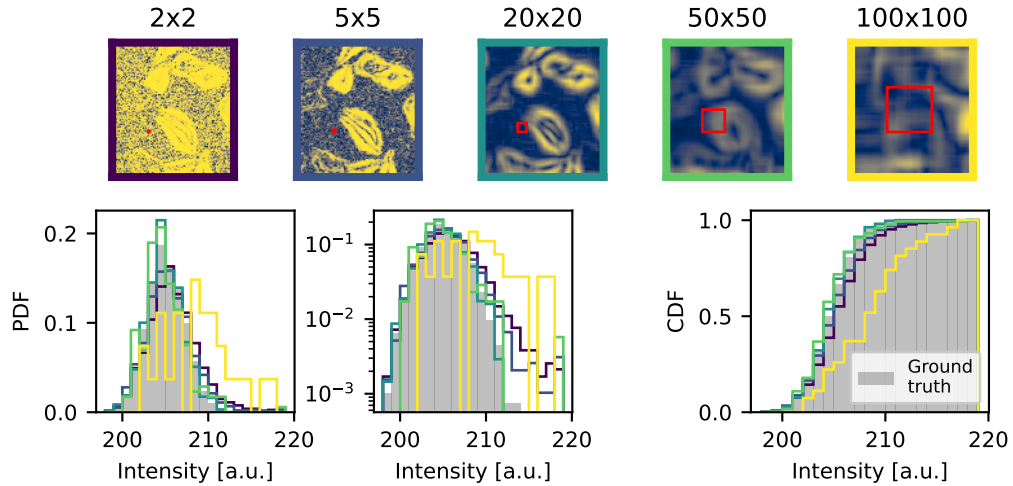


Fig. S2. The SMO is robust to changes in averaging kernel size. Top row: SMO images for different kernel sizes, which are shown to scale as red squares. Bottom row: background distributions estimated for each kernel size, compared to a manually selected region as ground truth.

3. A PREVIOUS SMOOTHING FILTER IMPROVES THE PERFORMANCE OF SMO

The SMO benefits from applying a previous smoothing filter to the intensity image, as it can better distinguish the intensity gradient from noise. Note that the thresholded mask obtained from the SMO image is applied to the original intensity image, and not the smoothed one.

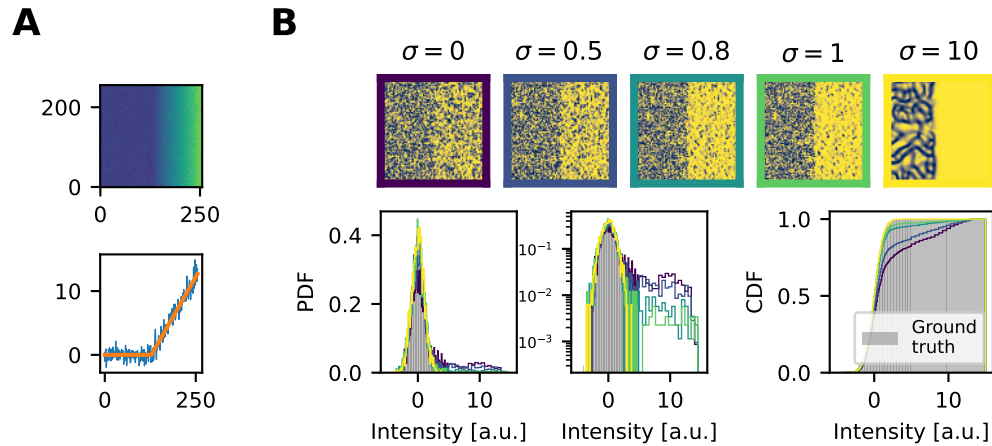


Fig. S3. (A) Intensity image simulated as a constant region (background) followed by a linearly increasing region (foreground), plus standard normal noise. (B) Top row: SMO images after applying a Gaussian filter of size σ . Bottom row: background distributions estimated, compared to the ground truth (left side of the image).

4. THE ROLLING BALL METHOD RECOVERS A BIASED BACKGROUND DISTRIBUTION

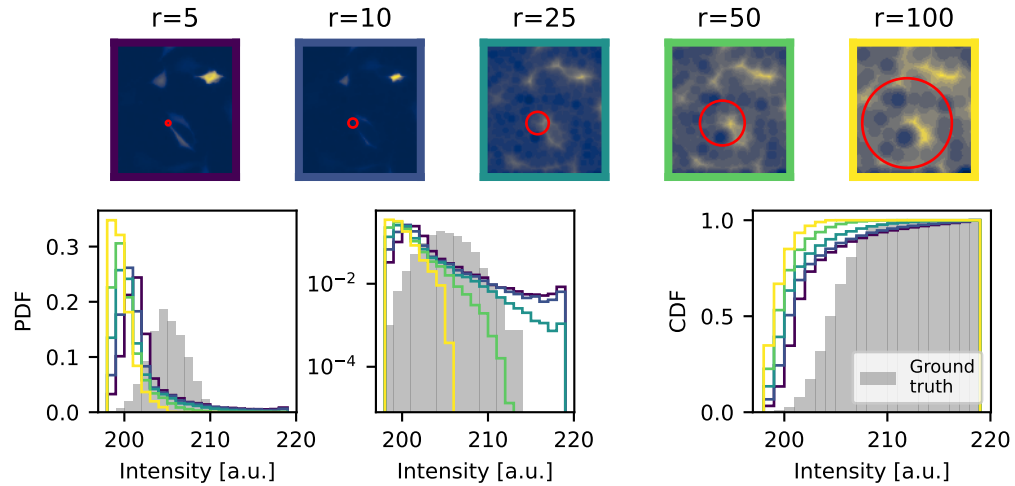


Fig. S4. The rolling ball method applied to the intensity image of figure 5. Top row: background estimation different radii r . Bottom row: background distributions estimated, compared to a manually selected region as ground truth.

5. SMO VALUES ARE INDEPENDENT FROM INTENSITY VALUES FOR ALL DISTRIBUTIONS

A uniform copula shows independence between intensity and SMO values for the different noise distributions.

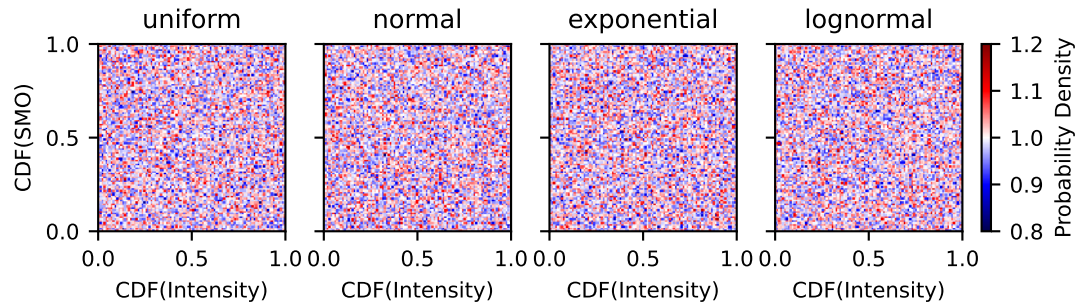


Fig. S5. Copulas of the joint distribution of intensity and SMO values for different noise distributions.

6. REFINEMENT OF THE SMO MASK WITH MORPHOLOGICAL OPERATIONS

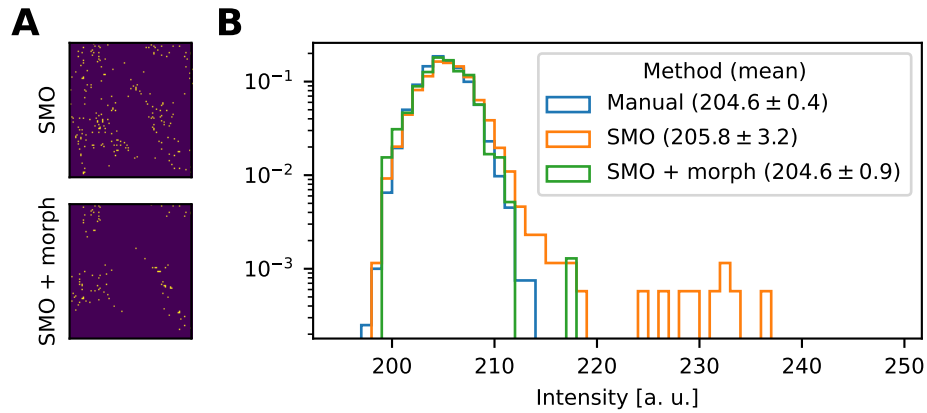


Fig. S6. Refining the SMO mask with morphological operations improves the background estimation. (A) Top: Original SMO mask. Bottom: refined with morphological operations to exclude points near to cell borders. (B) Background distributions estimated, compared to a manually selected region as ground truth.

7. CORRELATION BETWEEN BACKGROUND TO FOREGROUND AREA RATIO AND METHOD FAILURE IN BBBC025

When the number of background pixels is small compared to the foreground ones, all methods fail, estimating a higher value for the median background (Figure S7). While there is not much difference in the Hoechst channel, in the Mito, ERSyto and ERSytoBleed, all methods show a lower breaking point than the SMO method.

As we had no segmentation available as a ground truth, we used a manually selected threshold per channel to segment into foreground and background, and calculate the area ratio. The chosen thresholds were:

Channel	Threshold
Hoechst	540
PhGolgi	3500
Mito	1300
ERSyto	1100
ERSytoBleed	2000

Note that it is not the true ratio between foreground and background, but a proxy of it. In Figure S8, we showed a set of images corresponding to high or low values of this proxy, suggesting that it is measuring correctly the area ratio.

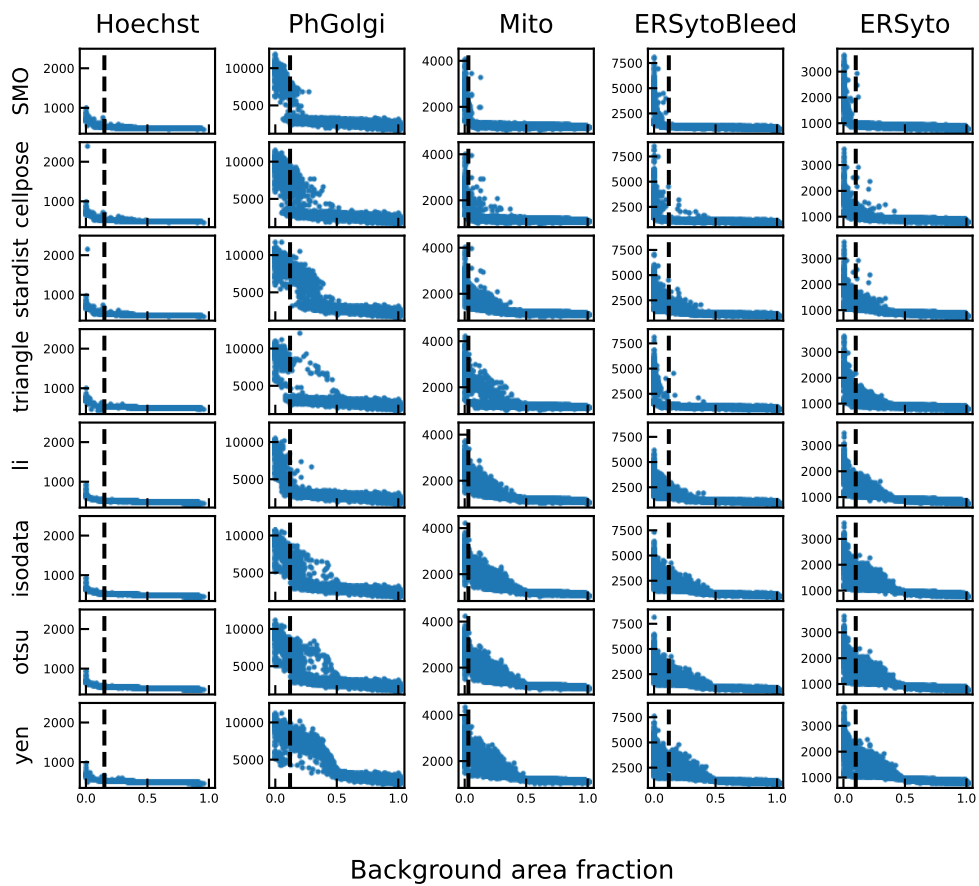


Fig. S7. Correlation between foreground-to-background area ratio and bias in median background intensity. The vertical dashed line marks the breaking point for the SMO method.

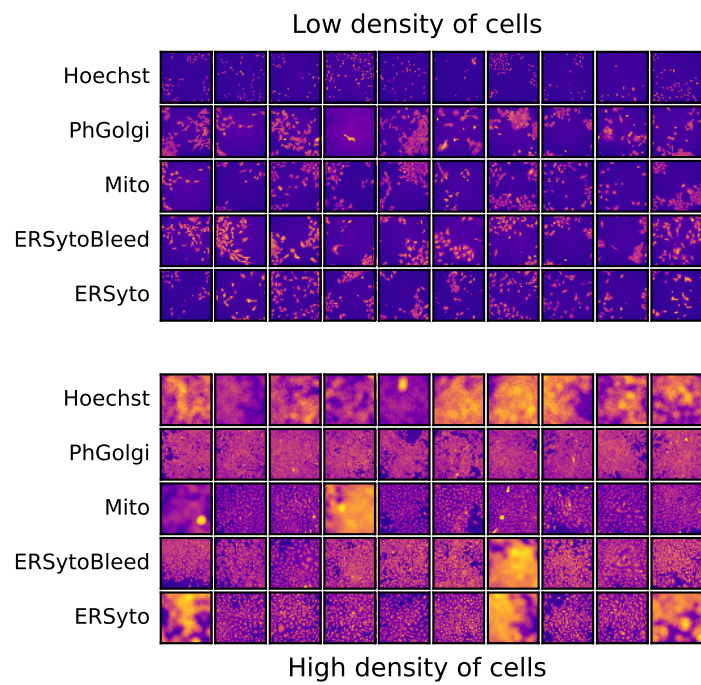


Fig. S8. Sample images of low and high density of cells. Top: sample images selected such that all methods reported a median value between the 10 and 20 percentile of the median background distribution. Bottom: all methods reported median background greater than the manually-selected thresholds.