

RAPID: Real-time image-based autofocus for all wide-field optical microscopy systems

L. Silvestri^{1,2,*}, M. C. Müllenbroich^{1,2}, I. Costantini², A. P. Di Giovanna², L. Sacconi^{1,2}, and F. S. Pavone^{2,3,1}

¹National Institute of Optics, National Research Council, Sesto Fiorentino, Italy

²European Laboratory for Non-linear Spectroscopy (LENS), Sesto Fiorentino, Italy

³Department of Physics and Astronomy, University of Florence, Italy

*email: silvestri@lens.unifi.it

Supplementary Material

Supplementary Notes.....	<i>pages 2-8</i>
Supplementary Methods.....	<i>pages 9-11</i>
Supplementary Tables.....	<i>pages 12-14</i>
Supplementary Figures.....	<i>pages 15-20</i>
Supplementary Videos Legend.....	<i>page 21</i>
References.....	<i>pages 22-23</i>

Supplementary Note 1

Commonly used autofocus methods in optical microscopy

Contrast-based methods

In these autofocus (AF) approaches, the focus position of the system is changed to explore the focus/contrast curve (Supplementary Fig. 1a), and find its maximum, which defines the ‘in-focus’ position. The main advantage of contrast-based AF is that it is purely image-based, and therefore can be applied to all wide-field microscopy systems. Conversely, the main disadvantage is that several images need to be collected to determine the optimum focal position. This implies that the system must stop acquiring ‘useful’ images when autofocusing, posing a strong limit on the applicability of these methods when the focus is rapidly changing (e.g. when tracking fast-moving samples), and in general increasing imaging time.

The two main components of contrast-based AF are the function(s) used to compute image contrast, and the strategy employed to find the peak of the focus/contrast curve. A variety of contrast functions have been described in the literature¹⁻⁸, which are extensively and thoroughly reviewed by in refs. 9,10. Given this plethora of functions, the user has to select the one(s) most appropriate to her samples and imaging system, finding the optimal trade-off between accuracy, reliability and computation time. Additionally, image pre-processing can be applied before computation of the proper contrast function. The ideal contrast function is characterized by i) a single, well-defined peak, ii) slow decay to detect a contrast gradient also away from focus, iii) noise insensitivity, iv) brightness insensitivity, v) resolution insensitivity. When the suitable contrast function is used, contrast-based AF can provide accurate focus stabilization within the microscope depth of focus (DOF), for several hours¹⁰. The choice of a good function allows robust operation also starting from a defocus of few tens of DOF¹¹.

Concerning the best strategy to find the optimal focus, several approaches have been described in the literature to limit the number of image samples needed, ranging from intelligent search¹² to parabolic fitting¹³. With this latter method it is possible to find the best focus with just 3 image samples; however it works only if the focus/contrast curve has a quadratic (or inverse quadratic) shape, which is not usually the case. More generally, the number of images used in practice by different implementations is of the order of ten¹⁰⁻¹².

Triangulation-based methods

In this class of AF techniques, an oblique light beam is projected onto the sample and reflected back by some fiducial surface located between the objective and the sample itself. As this surface moves back and forth with respect to the objective, the reflected light beam moves laterally (Supplementary Fig. 1b). By detecting this lateral shift, it is therefore possible to infer the axial position of the reflective surface with high accuracy, also well below the microscope DOF¹⁴. The operating range is dependent on the inclination of the light beam and on the aperture of the objective, and in general can span several tens of DOF¹⁵.

The crucial advantage of these methods is that they operate in real time, without having to suspend image acquisition. They are therefore well suited for the detection of fast defocus events. Furthermore, as the autofocus is completely independent from the imaging, these methods can perform equally well with high or low contrast. By contrast, the prominent weakness of triangulation-based strategies is that they read out the position of a fiducial reflective surface, which is used as a proxy for the position of the sample itself. Thus, the operation is correct only if the distance between the sample and the fiducial surface is fixed in time and space. Finally, these methods are completely not applicable to light sheet microscopy, where no flat reflective surface exist at all.

The typical merits and specification of both contrast-based and triangulation-based AF are summarized in Supplementary Table 1, together with those of RAPID.

Supplementary Note 2

Previous implementations of phase detection in optical microscopy

To the best of our knowledge, phase detection (PD) has been applied in microscopy only in two applications, i.e. 3D single-molecule localization and whole-slide histological imaging. In this Note we briefly recapitulate these two implementations, clarifying their specificity and limitations.

3D single-molecule localization

Yajima et al. described first the possibility of using PD to infer the axial position of single emitters¹⁶. This technique was then refined by Sun et al. in a method named ‘Parallax’¹⁷. In both approaches single-molecule detection algorithms (based on fitting of the point-spread-function) are performed separately in two images obtained from two distinct portions of the pupil. The localized positions of the same bead obtained from the two images are then compared to infer the axial position.

The main merits and limitations of these approaches stem from the use of single-molecule detection as first processing step. Indeed, as the lateral distance between the two images of the same emitter can be determined with sub-wavelength resolution, the accuracy in focus determination is quite high (almost 1/100 of microscope DOF¹⁷). Conversely, these methods are obviously applicable only to single-molecule settings and not to general microscopy images, where individual emitters are not distinguishable.

Whole-slide histological imaging

PD autofocus in whole-slide histological imaging has been recently reported^{18,19}. In the ‘Instantscope’ by Guo et al. two auxiliary camera are placed on the two oculars of a standard microscope, and two out-of-axis pinholes are used to select different regions of the pupil¹⁸. Phase correlation²⁰ is then used to compute mutual distance between the images. In the setup described by Liao et al., a dual-pinhole aperture is used to create two superimposed phase-shifted images on a single camera¹⁹. The secondary peaks of autocorrelation are then used to obtain the distance between the two superimposed images.

Both approaches can reach accuracy within DOF on a quite large range (orders of tens of DOF). However, the use of pinholes to isolate selected portions of the pupil limits the use of these methods only in

high light settings. Indeed, Liao et al. reported that the system can be applied to whole-slide fluorescence imaging only by using transmitted excitation light for AF purposes; fluorescence light itself is not enough brilliant for the PD system described there¹⁹.

The merits and specifications of the PD autofocus systems described in this Note are summarized in Supplementary Table 2, together with those of RAPID.

Supplementary Note 3

Theoretical and experimental RAPID accuracy and operating range

RAPID maximum theoretical accuracy

The resolution of RAPID, i.e. the minimum difference in defocus that can be detected, is related via Eq. (4) of Main Methods to the accuracy of the method in measuring d . From information theory principles, it is possible to derive the maximum registration accuracy that *whatever* registration method can achieve^{21,22}:

$$\delta d = \frac{\sigma}{\sqrt{\sum_{m,n} I_x^2(m,n)}} \quad (1)$$

where σ is the standard deviation of image noise, I_x is the spatial derivative of the image along the direction x , and m,n the pixel indices. Given the limited bandwidth of the microscope, the upper bound of $I_x(m,n)$ is given by $k_{\max} I(m,n)$, k_{\max} being the maximum spatial wave number allowed by the system, and $I(m,n)$ the pixel value at position (m,n) . Therefore, the lower bound for registration accuracy is given by:

$$\delta d \geq \frac{\sigma}{k_{\max} \langle I \rangle \sqrt{N}} = \frac{1}{k_{\max}} \frac{1}{\sqrt{N}} \frac{\sigma}{\langle I \rangle} = \frac{r_{\min}}{\sqrt{N} SNR} \quad (2)$$

with N being the number of pixels in the image, $\langle I \rangle$ the average image intensity, SNR the signal-to-noise ratio and r_{\min} the minimum distance resolved by the apparatus. As evident from Eq. (2), the theoretical best registration accuracy can be much higher than the microscope resolution, even by several orders of magnitude, as soon as N becomes a large number. From the above formula, it is also clear that image registration (and thus defocus measurement) can be very accurate quite away from the focus: even if r_{\min} becomes large (as it happens in defocused images), the denominator of Eq. (2) keeps the overall fraction small. For instance, if $N = 100 \times 100$ pixels and SNR is just 4 or 5, δd lower bound is in the submicron range even with $r_{\min} = 100 \mu\text{m}$, i.e. very away from the focus.

Experimental RAPID accuracy and operating range

In practice, d accuracy is limited by several factors, including pixel size, image contrast, signal-to-noise ratio, signal-to-background ratio, aberrations, etc. We empirically determined RAPID accuracy by estimating the root-mean-square (RMS) error in the detected defocus with respect to the best linear fit, over a

range 15-20 bigger than the objective depth of focus (Supplementary Table 2). On average, RAPID accuracy is about 70% of the objective depth of focus, providing enough accuracy to keep in focus microscopes with diffraction-limited resolution. With further improvements in the image registration pipeline (e.g. the use of methods providing sub-pixel accuracy), RAPID could also be applied in super-resolution settings, both to stabilize focus and provide a super-resolved readout of axial position.

The same factor limiting RAPID accuracy also pose practical bounds to its applicability range. Empirically, we observed reliable focus discrimination over a range 70 times larger than the objective depth of focus (Supplementary Fig. 5).

Regarding the applicability of the method in low-light settings, RAPID has been successfully applied to *in vivo* fluorescence imaging (Main text Fig. 1h, Supplementary Fig. 4), where images were characterized by signal-to-noise and signal-to-background ratios of 4.53 ± 0.07 and 1.66 ± 0.05 , respectively (mean \pm s.e.m., $N = 10$ images, $n = 5$ measurements per image). To quantify these values, we proceeded as follows: In the images collected by the auxiliary camera, we quantified ‘signal’ as the average counts from the 5 brighter cells in the field of view. We estimated the variance of total noise (in e^-) as the variance of shot noise (equal to the number of photoelectrons collected) plus the variance of read noise ($8 e^-$). Background was quantified as the average counts value in regions without fluorescent cells.

Supplementary Note 4

Imaging throughput in light-sheet microscopy with contrast-based AF

Contrast-based AF has been recently described in two different light-sheet microscopy (LSM) applications, i.e. long-term imaging of developing larvae¹⁰ and high-resolution reconstructions of entire murine brains²³. In both cases, AF is used to correct any possible sample-induced mismatch between the illumination light sheet and the focal plane of detection optics. The main difference (at least from an AF perspective) between the two applications is that in the former case the same volume is imaged at different time points, whilst in the latter different tiles are imaged sequentially. This is quite important since one can rely on correlations between different image acquisitions in order to reduce the amount of data collected for AF purposes.

For long-term *in vivo* imaging, one can leverage on the fact that defocus is slowly varying in time (at most about $0.2 \mu\text{m}/\text{min}$ ¹⁰) to avoid the complete refocusing of all the volume at each acquisition. Royer et al. reported that only 5% of the time is devoted to autofocus in standard conditions¹⁰.

Conversely, the spatial correlations between different tiles in high-resolution LSM of large specimen is not so stable, as shown in Supplementary Figure 6b. For this reason, in the protocol described by Tomer and colleagues the AF procedure is repeated in each single tile and at several depths, usually spaced by 1 mm²³. To estimate the amount of time t_{AF} spent for autofocusing, compared to ‘true’ imaging time t_{IM} (which is not reported in Ref. 23), we proceed as follows. We assume that each AF optimization takes 2 seconds, a time that includes image acquisition, computation of the focus/contrast curve and movement of the stage to the next point. We also assume that the camera acquire images at 50 Hz (the maximum speed of most sCMOS cameras in light-sheet mode). Then, for a stack of length l , we have:

$$\frac{t_{AF}}{t_{IM}} = \left(2 \frac{l}{d}\right) \left(\frac{\delta}{l} \frac{1}{0.02}\right) = \frac{100 \delta}{d} \quad (3)$$

where d is the step between different AF optimization depths, δ is the z-step of stack acquisition. The first parenthesis is t_{AF} , whilst the second is the inverse of t_{IM} . For typical values of d (1 mm) and δ ($2 \mu\text{m}$), we get $t_{AF}/t_{IM} = 20 \%$, while with a bigger (albeit commonly used) δ ($5 \mu\text{m}$) the ratio can be as high as 50 %.

SUPPLEMENTARY METHODS

Pupil-split images registration pipeline

One of the characteristic features of RAPID with respect to other phase-detection autofocusing systems, is that it works by computing the phase shift between the pupil-split images *globally*. In other words, it tries to register one images onto the other rigidly, rather than computing the phase displacement in small regions. In this perspective, it is crucial to obtain a reliable operation of the registration pipeline, robust to possible artefacts introduced by poor image quality, optical aberrations and so on. The registration pipeline used is described here in its three main subparts. All the software used for RAPID operation is freely available on https://github.com/ludovicosilvestri/RAPID_CLSM.

1. Pupil-split images quality check

Before computing the mutual registration between the pupil-split images, a quality check step is performed to skip the registration when the information content of the images is too low, e.g. because the microscope field of view is momentarily outside of the specimen, or because the image is highly saturated. This is necessary to avoid wrong focusing operation that may render RAPID system instable. To this aim, three indicators of quality are considered: average value, standard variation (the ratio between standard deviation and average value) and the percentage of saturating pixels. If all three parameters are above a user-defined threshold, images successfully pass the quality check. Otherwise, the images are discarded and no registration is computed.

2. Pupil-split images pre-processing

The mutual displacement of the two pupil-split images is obtained by finding the peak of cross-correlation between the images themselves. As discussed in the Supplementary Note, the accuracy of this procedure is limited by image quality and contrast. To achieve a more reliable and accurate RAPID operation, images are pre-processed to obtain a more pronounced peak of the cross-correlation map. We employed up to three different pre-processing strategies depending on the sample under observation (Supplementary Table 1). Optionally, image lookup table was inverted before pre-processing.

a. Image binarization

Images are first converted to unsigned 8-bit type. A binarization threshold is then computed using Ridler's clustering method²⁴: the threshold value T is the one satisfying the condition $T = \mu_1 + \mu_2/2$, where μ_1 and μ_2 are the average values of all the pixel values above and below T , respectively. Images are then binarized by putting all the pixels with value below T to 0, and all the ones with values above T to 255.

b. Difference of Gaussians

Images are first converted to 32-bit floating point type. Then, two distinct copies of the images are blurred by convolution with two Gaussian kernels with different sigma. The 'sharper' image (the one convoluted with the Gaussian kernel with smaller sigma) is subsequently subtracted from the 'more blurred' one. This procedure allows enhancing features of the original images²⁵. Finally, the difference image is rescaled to the range [0 255] and converted to unsigned integer 8-bit type.

c. Edges enhancement

Images are first converted to unsigned integer 8-bit type. Then, a Canny edge detection filter is applied²⁶, providing a binary image with bright lines located at the edges of the original one. Afterwards, a smoothing filter is applied, followed by a morphological dilation, to smooth possible differences between the two images and boost the efficacy of cross-correlation.

3. Proper image registration and final quality check

Pupil-split images, pre-processed according to one of the methods described above, are registered by finding the maximum of normalized cross-correlation (NCC). If more than one pre-processing method was used, NCC maxima obtained with different pre-processing steps are combined by averaging.

A final quality check is performed on the NCC results to help discarding wrong registration outcomes. Computed phase between images is discarded if it exceeds a user-defined threshold; additionally, when several pre-processing strategies are used, phase is discarded if the NCC maxima locations calculated with different pre-processing are too far away. If the quality check is passed, the phase between pupil-split images is sent to

a proportional-integrative (PI) feedback system to correct microscope defocus. Otherwise, computation results are discarded.

Supplementary Table 1

<i>AF approach</i>	<i>Typical accuracy (DOF)</i>	<i>Typical range (DOF)</i>	<i># of extra images needed</i>	<i>Applicability</i>		
				<i>Quick focus changes</i>	<i>3D sample environments</i>	<i>Light-sheet microscopy</i>
Contrast maximization ¹⁰⁻¹³	0.2 to 2	10 to 100	~ 10	No	Yes	Yes
Triangulation ^{14,15}	0.05 to 0.5	10 to 300	0	Yes	No	No
<i>RAPID</i>	~ 0.7	~ 70	0	Yes	Yes	Yes

Comparison between the typical features of contrast-based AF, triangulation methods, and RAPID.

Supplementary Table 2

<i>Method</i>	<i>Accuracy (DOF)</i>	<i>Range (DOF)</i>	<i>Demonstrated time stability</i>	<i>Min S/N</i>	<i>Min S/B</i>	<i>Fluorescence</i>	<i>Dense samples</i>
<i>Parallax</i> ¹⁷	~ 0.01	~ 8	minutes	~ 120	~1300	Yes	No
<i>Instantscope</i> ¹⁸	~ 0.3	~ 300	N/A	N/A	N/A	No	Yes
Single-camera <i>Instantscope</i> ¹⁹	~ 0.11	~ 25	N/A	N/A	N/A	No	Yes
<i>RAPID</i>	~ 0.7	~ 70	> 12 hours	~ 4.5	~ 1.7	Yes	Yes

Comparison between the features and limitations of RAPID and of previous implementations of phase detection in optical microscopy.

Supplementary Table 3

<i>Objective</i>	<i>Illumination mode</i>	<i>Objective DOF*</i> (μm)	<i>RMS of residuals</i> (μm)	<i>Fit range</i> (μm)	<i>RMS/DOF ratio</i>
Nikon Plan Apo 60X, NA 1.4, Oil immersion	Brightfield	0.39	0.35	6	0.90
	Epifluorescence		0.26		0.65
Nikon Plan Fluor 40X, NA 0.6	Brightfield	1.39	0.84	20	0.61
	Epifluorescence		0.42		0.30
Olympus XLPLN10XSVMP 10X, NA 0.6, mult. imm.	Light-sheet	2.00	2.08	40	1.04
				<i>Average</i> \pm <i>S.E.M.</i>	0.70 \pm 0.14

RAPID accuracy with different objectives used.

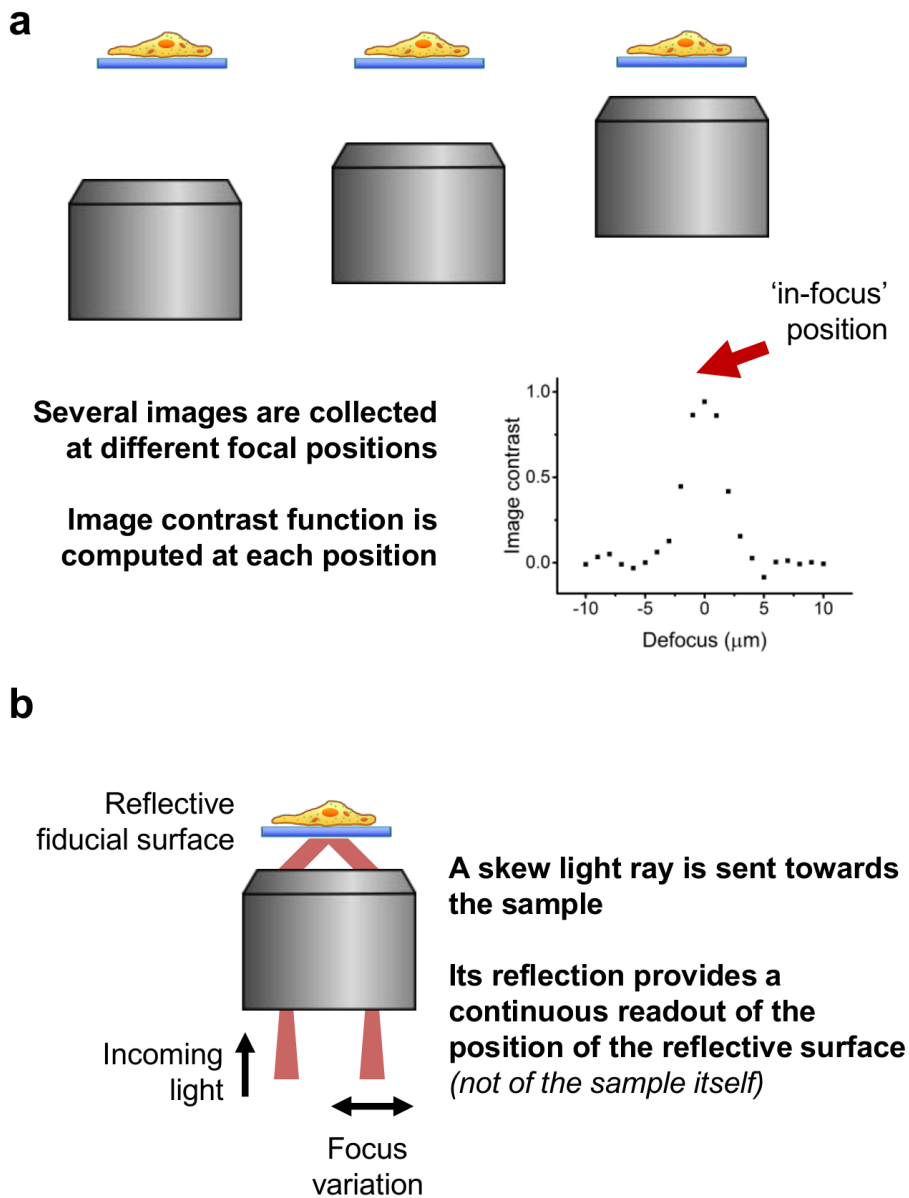
*calculated using the formula $DOF = n\lambda/NA^2$, with wavelength $\lambda = 0.5 \mu\text{m}$, refractive index $n = 1.52$ for the 60X objective, $n = 1$ for the 40X, and $n = 1.46$ for the 10X.

Supplementary Table 4

<i>Experiment</i>	<i>Objective used</i>	<i>Illumination mode</i>	<i>Pre-processing before image registration</i>
Whole-slide histological imaging	Nikon Plan Fluor 40×, NA 0.6	Brightfield	Binarization and Difference of Gaussians
Long-term <i>S. Cerevisiae</i> imaging	Nikon Plan Apo 60×, NA 1.4, Oil immersion	Brightfield	Binarization on inverted image
Long-term <i>S. Pombe</i> imaging	Nikon Plan Apo 60×, NA 1.4, Oil immersion	Epifluorescence, excitation with LED @ 470 nm	Binarization
<i>C. Elegans</i> 3D tracking	Nikon Plan Fluor 10×, NA 0.3	Brightfield	Edges enhancement
<i>Thy1-GFP-M</i> mouse brain imaging	Olympus XLPLN10XSVM 10×, NA 0.6, multiple immersions	Light-sheet, excitation with laser @ 491 nm	Binarization and Difference of Gaussians
Mouse brain vasculature imaging	Olympus XLPLN10XSVM 10×, NA 0.6, multiple immersions	Light-sheet, excitation with laser @ 561 nm	Binarization and Difference of Gaussians
Mouse brain imaging of cell nuclei	Olympus XLPLN10XSVM 10×, NA 0.6, multiple immersions	Light-sheet, excitation with laser @ 561 nm	Binarization and Difference of Gaussians

Objectives and image pre-processing used in the various experiments reported in the paper.

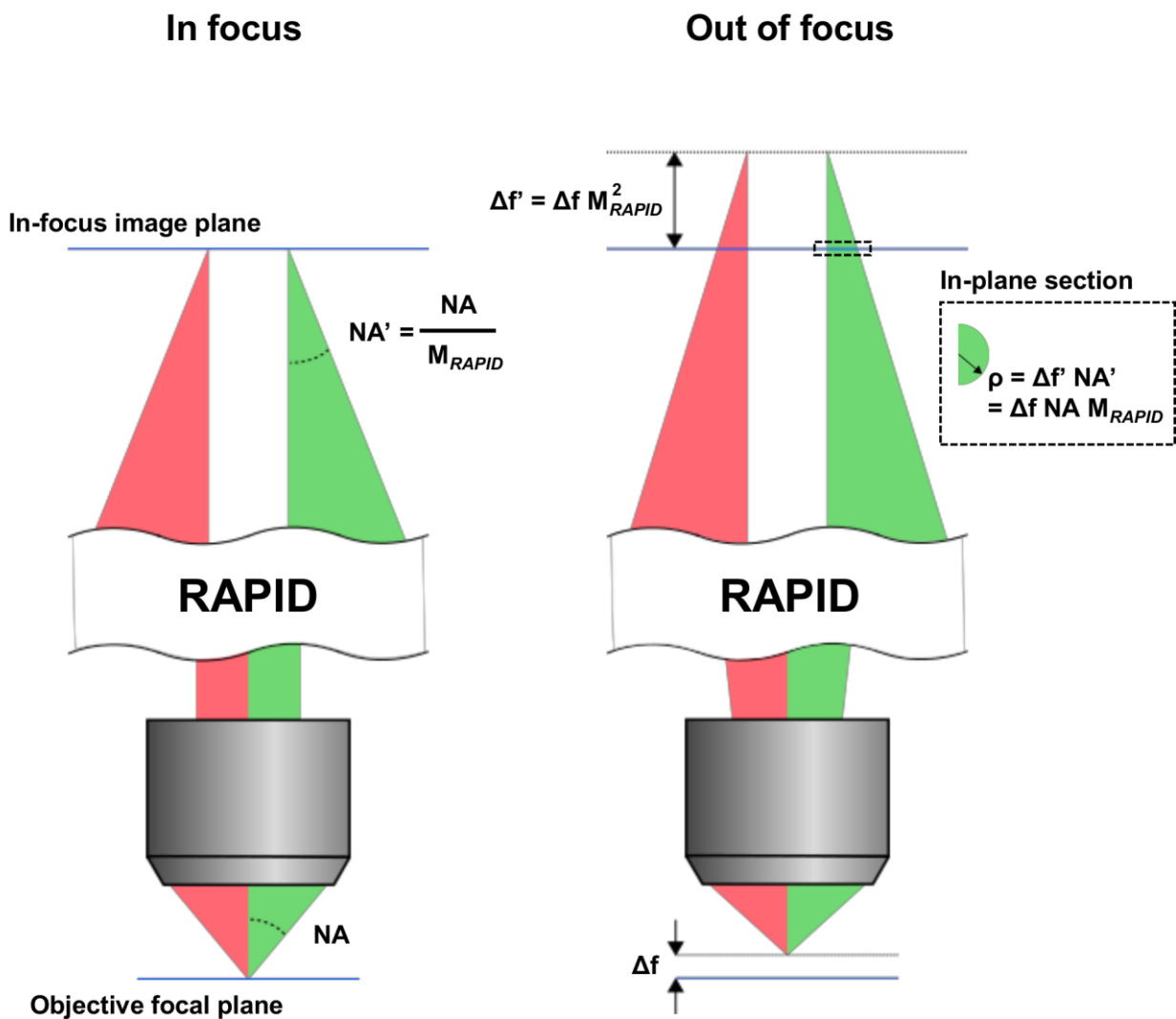
Supplementary Figures



Supplementary Figure 1. The two main approaches used in microscopy for defocus correction.

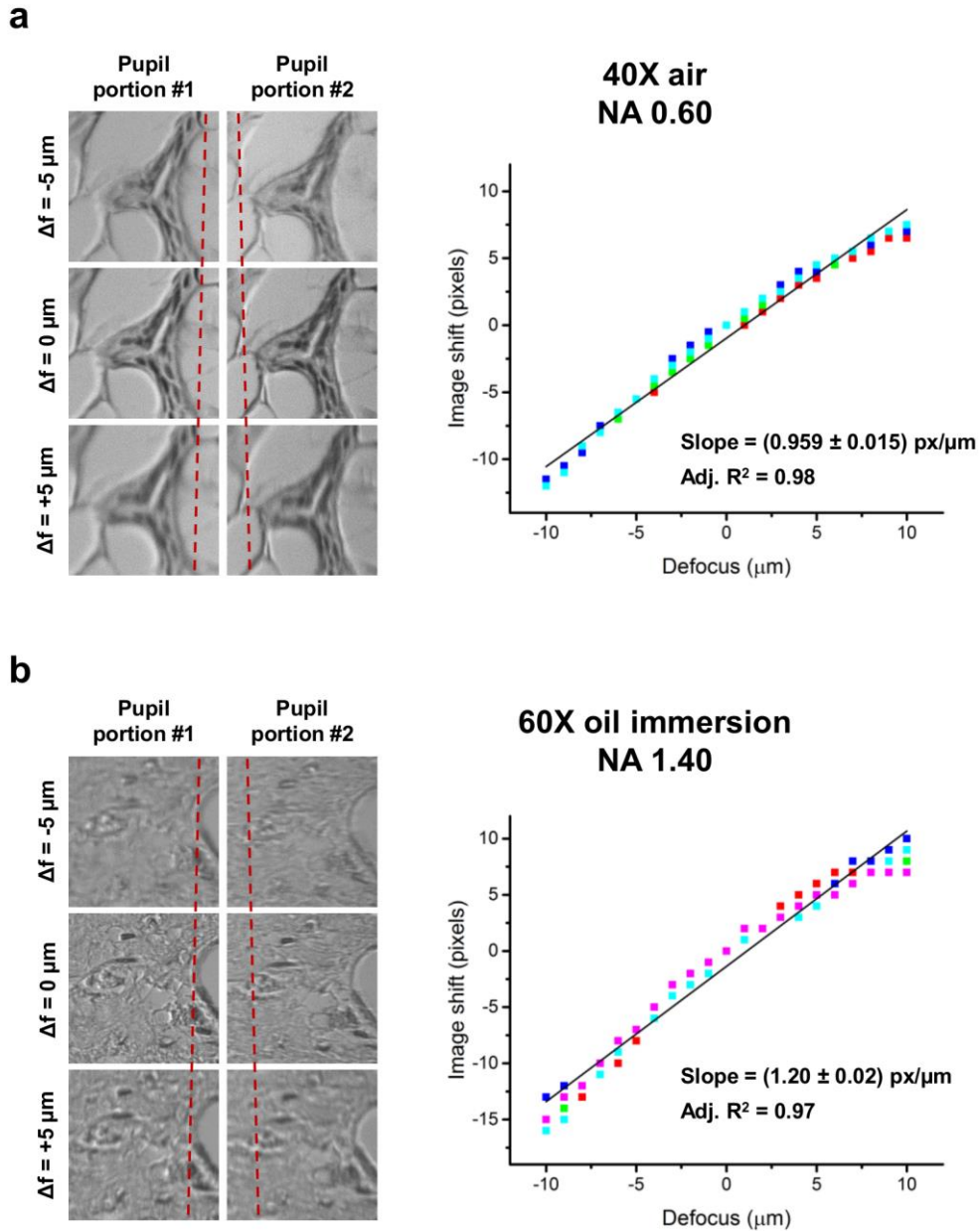
In contrast-based methods (a), different images are acquired at various focal positions by moving either the objective or the sample, or the sheet of light in light-sheet microscopy. Afterwards, a contrast function is used to evaluate image sharpness as a function of defocus. The position corresponding to the contrast peak is considered ‘in-focus’.

In triangulation methods (b), a skewed ray of light is projected onto the sample, and it is reflected by some fiducial surface (e.g. the coverslip). Vertical motion of this surface results in lateral displacement of the reflected beam. By assuming a fixed distance between the sample and the fiducial surface, this displacement can be used to infer microscope defocus.



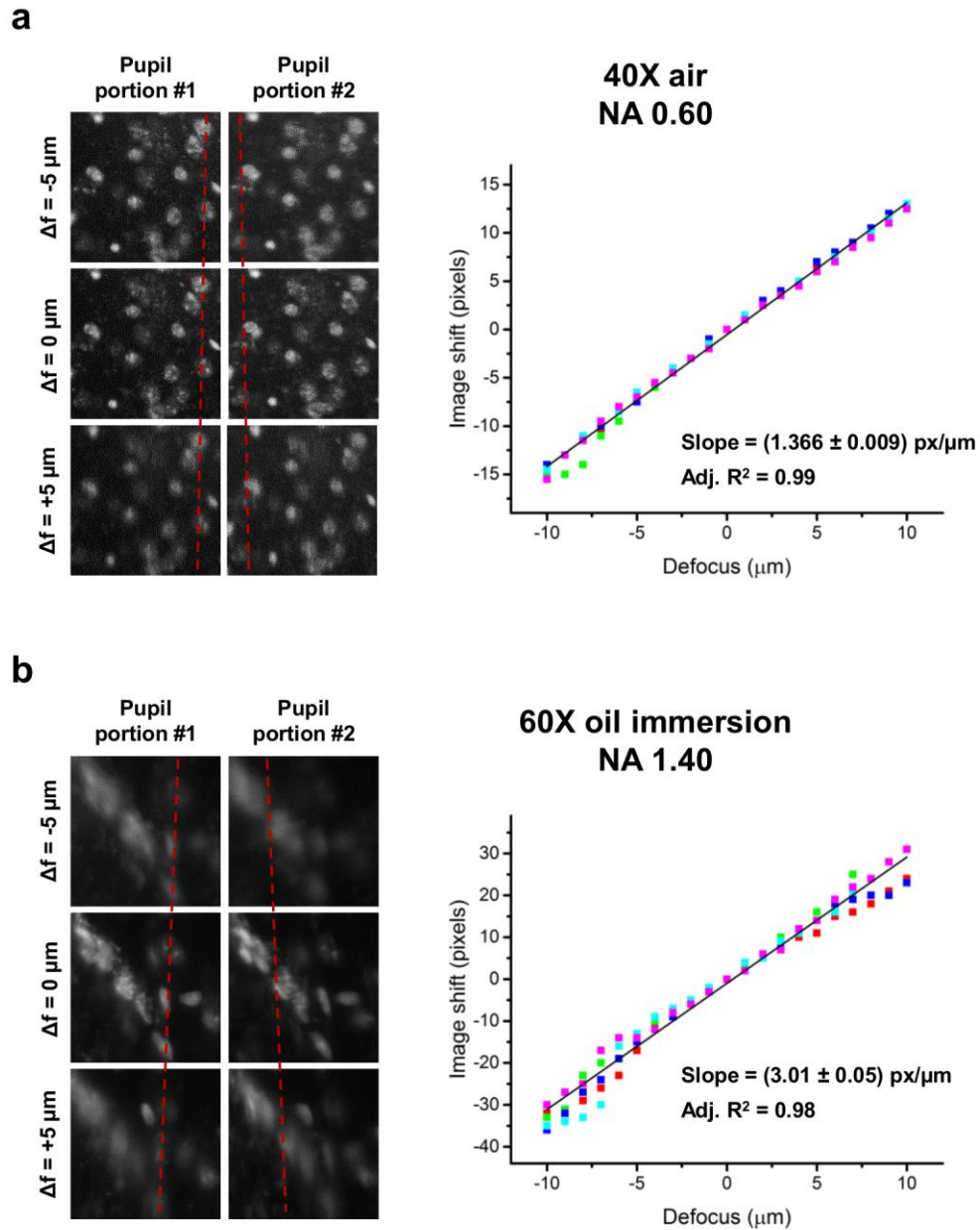
Supplementary Figure 2.

Relation between mutual image displacement and defocus in RAPID. The radius ρ of the half-circular bundle of ray intersecting the in-focus image plane depends linearly on the defocus Δf via the formula indicated in the image. The two portions of the pupil splitted by the RAPID system are shown in red and green. Only the microscope objective and the image space of RAPID (where the phase detection camera is placed) are shown.



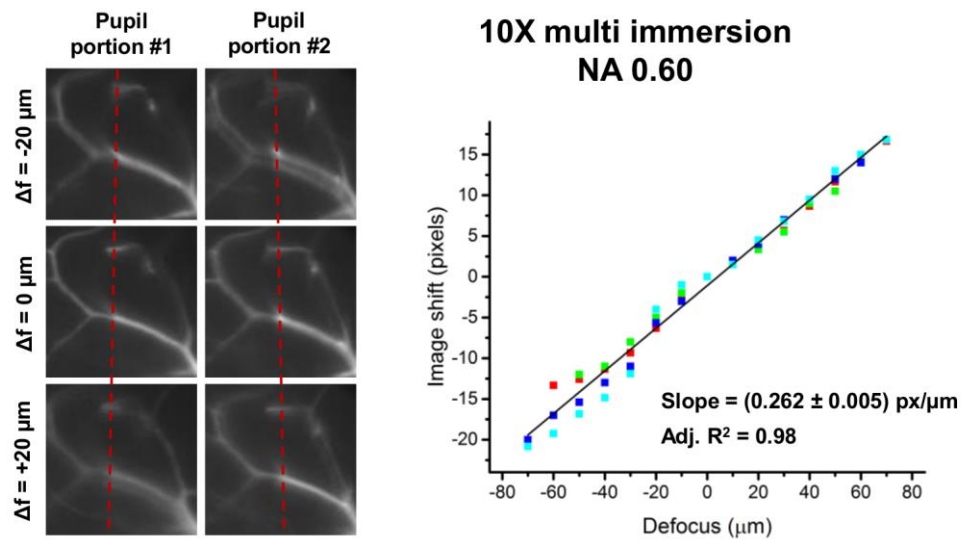
Supplementary Figure 3.

Mutual lateral displacement of the two pupil-split images in brightfield illumination settings, for two different objectives (a and b). For these measurements, histological keloid samples stained with hematoxylin and eosin were used. Left panels show examples of the image shifts, with a dashed red line to help the reader visualizing the small lateral displacements. Right panels show image shifts determined by automatic image analysis as a function of defocus; different colors identify distinct defocus stacks. A linear fit with the measured slope and a measure of fit goodness (the adjusted R^2 score) is also shown.



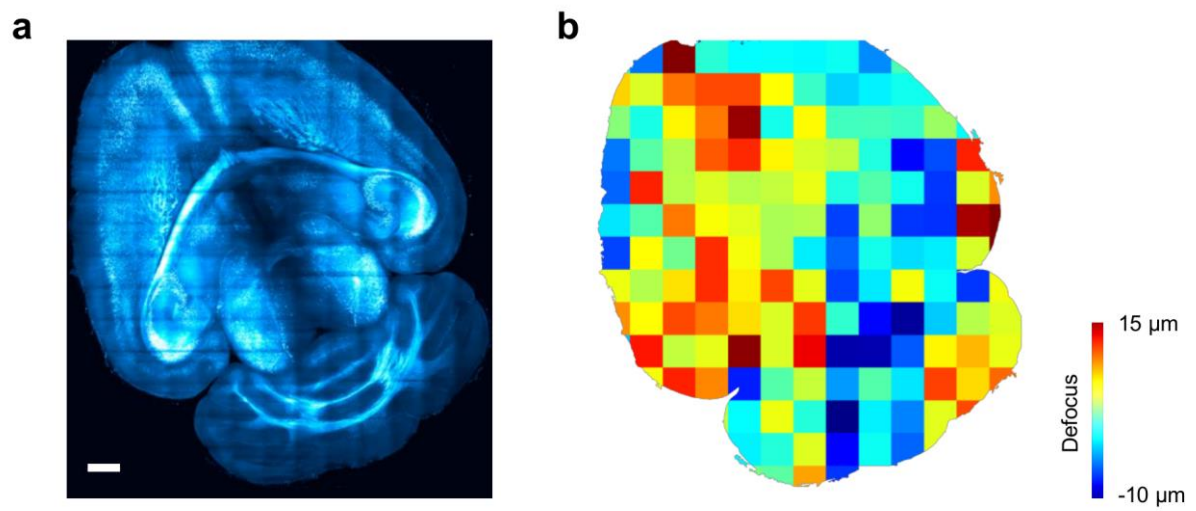
Supplementary Figure 4.

Mutual lateral displacement of the two pupil-split images in epifluorescence illumination settings, for two different objectives (a and b). For these measurements, mouse brain slices stained with SYTOX Green Nucleic Acid Stain (ThermoFisher) were used. Left panels show examples of the image shifts, with a dashed red line to help the reader visualizing the small lateral displacements. Right panels show image shifts determined by automatic image analysis as a function of defocus; different colors identify distinct defocus stacks. A linear fit with the measured slope and a measure of fit goodness (the adjusted R^2 score) is also shown.



Supplementary Figure 5.

Mutual lateral displacement of the two pupil-split images in light-sheet illumination settings. For these measurements, cleared mouse brain with vasculature fluorescent filling were used. Left panel shows examples of the image shifts, with a dashed red line to help the reader visualizing the small lateral displacements. Right panel shows image shifts determined by automatic image analysis as a function of defocus; different colors identify distinct defocus stacks. A linear fit with the measured slope and a measure of fit goodness (the adjusted R^2 score) is also shown. Note that quite good linearity is observed over a range of $140 \mu\text{m}$, about 70 times larger than the objective depth of focus ($\approx 2 \mu\text{m}$).



Supplementary Figure 6.

Virtual slab (500 μm thick) from the brain of a thy1-GFP-M mouse (a). Scale bar, 1mm. The average defocus for each tile is shown in (b), denoting a clear focal difference between the two light sheets used to illuminate the two halves of the brain, as well as variable defocusing across tiles.

Supplementary Videos Legend

Supplementary Video 1

Mutual lateral motion of the two pupil-split images as a function of defocus. The amount of defocus is indicated in the upper left corner. A reference target (R1L3S2P, Thorlabs) was used to avoid any potential specimen-induced artefact.

Supplementary Video 2

Long-term brightfield imaging of *S. Cerevisiae* cell culture using RAPID (left) or with a fixed focal position (right). Elapsed time is indicated in the upper left corner as hours:minutes.

Supplementary Video 3

Long-term epifluorescence imaging of *S. Pombe* cell culture using RAPID (left) or with a fixed focal position (right). Elapsed time is indicated in the upper left corner as hours:minutes.

Supplementary Video 4

Tracking of freely moving *C. Elegans* in brightfield illumination, using RAPID (left) or with a fixed focal position (right). In the RAPID panel is also indicated the axial position of the stage, used to reconstruct the 3D trajectory of the worm. The movie is shown at double speed with respect to real time.

Supplementary Video 5

Image stack from a cleared mouse brain with vasculature stained, acquired with light sheet microscopy, using RAPID (left) or with a fixed focal position (right). The stack is the same shown in Fig. 2d (Main text). Imaging depth inside tissue is reported in the upper left corner.

References

- 1 Brenner, J. F. *et al.* An automated microscope for cytologic research a preliminary evaluation. *Journal of Histochemistry & Cytochemistry* **24**, 100-111 (1976).
- 2 Groen, F. C. A., Young, I. T. & Lighthart, G. A Comparison of Different Focus Functions for Use in Autofocus Algorithms. *Cytometry* **6**, 81-91, doi:DOI 10.1002/cyto.990060202 (1985).
- 3 Vollath, D. The Influence of the Scene Parameters and of Noise on the Behavior of Automatic Focusing Algorithms. *J Microsc-Oxford* **151**, 133-146 (1988).
- 4 Yeo, T. T. E., Ong, S. H., Jayasooriah & Sinniah, R. Autofocusing for Tissue Microscopy. *Image Vision Comput* **11**, 629-639, doi:Doi 10.1016/0262-8856(93)90059-P (1993).
- 5 Nayar, S. K. & Nakagawa, Y. Shape from Focus. *IEEE transactions on pattern analysis and machine intelligence* **16**, 824-831, doi:Doi 10.1109/34.308479 (1994).
- 6 Subbarao, M., Choi, T. & Nikzad, A. Focusing Techniques. *Opt Eng* **32**, 2824-2836, doi:Doi 10.1117/12.147706 (1993).
- 7 Kristan, M., Pers, J., Perse, M. & Kovacic, S. A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recogn Lett* **27**, 1431-1439, doi:10.1016/j.patrec.2006.01.016 (2006).
- 8 Xie, H., Rong, W. & Sun, L. Construction and evaluation of a wavelet-based focus measure for microscopy imaging. *Microsc Res Tech* **70**, 987-995, doi:10.1002/jemt.20506 (2007).
- 9 Sun, Y., Duthaler, S. & Nelson, B. J. Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microsc Res Tech* **65**, 139-149, doi:10.1002/jemt.20118 (2004).
- 10 Royer, L. A. *et al.* Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms. *Nat Biotechnol* **34**, 1267-1278, doi:10.1038/nbt.3708 (2016).
- 11 Geusebroek, J. M., Cornelissen, F., Smeulders, A. W. & Geerts, H. Robust autofocusing in microscopy. *Cytometry* **39**, 1-9 (2000).
- 12 Gu, C. C., Wu, K. J., Hu, J., Hao, C. & Guan, X. P. Region sampling for robust and rapid autofocus in microscope. *Microsc Res Tech* **78**, 382-390, doi:10.1002/jemt.22484 (2015).
- 13 Yazdanfar, S. *et al.* Simple and robust image-based autofocusing for digital microscopy. *Opt Express* **16**, 8670-8677 (2008).
- 14 Neumann, B., Damon, A., Hogenkamp, D., Beckmann, E. & Kollmann, J. A Laser-Autofocus for Automatic Microscopy and Metrology. *Sensor Actuator* **17**, 267-272, doi:Doi 10.1016/0250-6874(89)80090-3 (1989).
- 15 Li, Q. X., Bai, L. F., Xue, S. F. & Chen, L. Y. Autofocus system for microscope. *Opt Eng* **41**, 1289-1294, doi:10.1117/1.1473639 (2002).
- 16 Yajima, J., Mizutani, K. & Nishizaka, T. A torque component present in mitotic kinesin Eg5 revealed by three-dimensional tracking. *Nature structural & molecular biology* **15**, 1119-1121, doi:10.1038/nsmb.1491 (2008).
- 17 Sun, Y., McKenna, J. D., Murray, J. M., Ostap, E. M. & Goldman, Y. E. Parallax: high accuracy three-dimensional single molecule tracking using split images. *Nano letters* **9**, 2676-2682, doi:10.1021/nl901129j (2009).
- 18 Guo, K., Liao, J., Bian, Z., Heng, X. & Zheng, G. InstantScope: a low-cost whole slide imaging system with instant focal plane detection. *Biomedical optics express* **6**, 3210-3216, doi:10.1364/BOE.6.003210 (2015).
- 19 Liao, J. *et al.* Single-frame rapid autofocusing for brightfield and fluorescence whole slide imaging. *Biomedical optics express* **7**, 4763-4768, doi:10.1364/BOE.7.004763 (2016).
- 20 Foroosh, H., Zerubia, J. B. & Berthod, M. Extension of phase correlation to subpixel registration. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* **11**, 188-200, doi:10.1109/83.988953 (2002).
- 21 Robinson, D. & Milanfar, P. Fundamental performance limits in image registration. *IEEE Transactions on Image Processing* **13**, 1185-1199 (2004).
- 22 Pham, T. Q., van Vliet, L. J. & Schutte, K. in *Electronic Imaging 2005*. 169-180 (International Society for Optics and Photonics).
- 23 Tomer, R., Ye, L., Hsueh, B. & Deisseroth, K. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nature protocols* **9**, 1682-1697, doi:10.1038/nprot.2014.123 (2014).
- 24 Ridler, T. & Calvard, S. Picture thresholding using an iterative selection method. *IEEE transactions on Systems, Man and Cybernetics* **8**, 630-632 (1978).

- 25 Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision* **60**, 91-110 (2004).
- 26 Canny, J. A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence* **8**, 679-698 (1986).