

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

CIDRE: an illumination-correction method for optical microscopy

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Kevin Smith, Yunpeng Li, Filippo Piccinini, Gábor Csúcs, Csaba Balazs, Alessandro Bevilacqua, et al. (2015). CIDRE: an illumination-correction method for optical microscopy. NATURE METHODS, 12(5), 404-406 [10.1038/nmeth.3323].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/428970> since: 2021-03-11

*Published:*

DOI: <http://doi.org/10.1038/nmeth.3323>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

**Smith, K., Li, Y., Piccinini, F. *et al.* CIDRE: an illumination-correction method for optical microscopy. *Nat Methods* 12, 404–406 (2015).**

The final published version is available online at:  
<https://doi.org/10.1038/nmeth.3323>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

# CIDRE: an illumination-correction method for optical microscopy

Kevin Smith<sup>1</sup>, Yunpeng Li<sup>2</sup>, Filippo Piccinini<sup>3</sup>, Gabor Csucs<sup>1</sup>, Csaba Balazs<sup>1</sup>, Alessandro Bevilacqua<sup>3,4</sup> & Peter Horvath<sup>5-7</sup>

<sup>1</sup>Scientific Center for Optical and Electron Microscopy, ETH Zürich, Zurich, Switzerland. <sup>2</sup>Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. <sup>3</sup>Advanced Research Center on Electronic Systems, University of Bologna, Bologna, Italy. <sup>4</sup>Department of Computer Science and Engineering, University of Bologna, Bologna, Italy. <sup>5</sup>Institute of Biochemistry, ETH Zürich, Zurich, Switzerland. <sup>6</sup>Synthetic and System Biology Unit, Biological Research Center, Szeged, Hungary. <sup>7</sup>Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland. Correspondence should be addressed to K.S. ([kevin.smith@unibas.ch](mailto:kevin.smith@unibas.ch)).

RECEIVED 22 APRIL 2014; ACCEPTED 16 DECEMBER 2014; PUBLISHED ONLINE XX XXXXXX 2015; [DOI:10.1038/NMETH.XXXX](https://doi.org/10.1038/NMETH.XXXX)

**Uneven illumination affects every image acquired by a microscope. It is often overlooked, but it can introduce considerable [AU: Use of “significant” is reserved for the statistical sense; instances in the paper have been changed to “considerable” or “substantial.”] bias to image measurements. The most reliable correction methods require special reference images, and retrospective alternatives do not fully model the correction process. Our approach overcomes these issues for most optical microscopy applications without the need for how**

No optical system is ideal. Inhomogeneous illumination is present in every image acquired by a microscope. Many factors, including misaligned optics, dust, nonuniform light sources and vignetting, contribute to uneven illumination<sup>1</sup>.

It is increasingly common for light microscopes to be used as quantitative instruments even though seemingly minor shifts in illumination can corrupt measurements and invalidate subsequent analyses. For example, we found that uneven illumination increased the false detections and missed detections by CellProfiler<sup>2</sup> on images of yeast cells by 35% when illumination correction was neglected (**Supplementary Fig. 1c–f**). Other routine measurements can be affected as well. Uneven illumination substantially reduced the measurements of the mean intensity and mean area of GFP-stained HeLa cells in the corner of the image relative to the center (**Supplementary Fig. 1g–l**).

The consequences of ignoring uneven illumination are often underestimated, as reflected in our survey of microscope users (**Supplementary Note 1**). The magnitude of intensity loss attributed to vignetting, that is, falloff of intensity from the center of the image, is often substantially stronger than assumed. Data from 11 ordinary microscope setups revealed that between 10% and 40% less light is typically recorded at the dimmest region of the image (**Supplementary Note 2**). Intensity loss is even more severe for cameras with large sensor areas or wide apertures, such as scientific complementary metal-oxide semiconductor (sCMOS) devices, which can experience a falloff greater than 50% (**Supplementary Note 2**).

The most common approach for correcting uneven illumination reverses the image formation process, attempting to recover the true image,  $I$ , from the image observed by the sensor,  $I_0$ . Distortions to the observed image are modeled by a linear intensity gain function  $v$  and an additive term  $z$ ;  $I_0(x) = I(x)v(x) + z(x)$ , where  $I_0(x)$  is the intensity observed at location  $x$ . The intensity gain models attenuations to the signal (**Fig. 1a**). An additive or zero-light term models contributions present even if no light is incident on the sensor, mainly camera offset and fixed-

pattern thermal noise. It is usually nearly uniform, varying by only a few intensity values. The uncorrupted image is recovered by reversing the image formation process

$$I(x) = \frac{I_0(x) - z(x)}{v(x)} \quad (1)$$

Although simple at first glance, in practice  $v$  and  $z$  cannot be known exactly, which has prompted the development of a variety of correction methods (**Supplementary Note 3**). Prospective methods estimate the correction surfaces from special reference images collected during acquisition<sup>3,4</sup>, whereas retrospective methods rely on the actual image data<sup>5,6</sup>. Prospective methods are regarded as more reliable because they empirically estimate the terms in equation (1), whereas retrospective methods are more practical because they do not require special acquisitions.

We introduce a new retrospective method, corrected intensity distributions using regularized energy minimization (CIDRE), which achieves correction quality similar to that of prospective methods. Unlike existing retrospective methods, CIDRE estimates both  $v$  and  $z$ . The key insight to our approach is that the distribution of intensities from a single location across many images is related to an underlying distribution common to all locations by a linear transform. This assumes that objects may appear anywhere in the image with equal probability. Local-intensity distributions from a finite set of observed images are simply linear transforms of a sampling of the underlying distribution (**Fig. 1b**). The parameters of the transform can be visualized in a quantile-quantile plot, where  $v$  corresponds to the slope and  $b$  corresponds to the  $y$  intercept.

CIDRE estimates  $v$ ,  $b$  and  $z$  simultaneously for all locations by minimizing a regularized energy function composed of several terms (**Fig. 1c**). The first term is a robust regression that ensures  $v$  and  $b$  fit the data (**Fig. 1d**). The second term reduces noise and guarantees a smooth correction surface, encouraging neighboring distributions to agree on similar values for  $v$  (**Fig. 1e**). The third term estimates  $z$  by finding the common point where all regression lines intersect (**Fig. 1f**). Corrected images are obtained using equation (1) (**Fig. 1g**). A complete explanation is provided in **Supplementary Note 4**.

We compared our approach to 12 commonly used methods in a series of tests on 12 data sets. Eleven of these data sets represent typical microscopy setups with various microscopes, sample preparations, staining, light sources, magnifications and sensors (**Supplementary Table 1** and **Supplementary Note 2**). As a gold standard, we used prospective correction in which  $v$  is estimated as the average of several reference images or empty images and  $z$  is estimated as the average of several dark-frame images. To measure correction quality, we collected hundreds of pairs of overlapping images for each data set, precisely aligned them and reported the mean of absolute differences for each pair of corresponding pixels in the overlapping regions (**Supplementary Note 5**). The scores are normalized by the mean score of the uncorrected pairs (**Fig. 2**). Although this provides a reasonable estimate of the correction quality, unavoidable differences between image pairs may cast some doubt on our measure. To address this, we generated a twelfth data set of synthetic images and distorted them using a known model (**Supplementary Note 6**), allowing us to directly measure the disagreement between the true and corrected images.

Our tests highlight the inadequacies of current illumination-correction practices. Single-image retrospective methods generally made the illumination more uneven instead of correcting it. Multi-image retrospective methods showed some improvement, but no method consistently achieved satisfactory performance. The gold-standard prospective methods performed best among existing methods, but their reliance on reference images limits their usefulness.

CIDRE is a retrospective method and thus does not require calibration images; therefore, it can always be applied to previously collected data. Unlike other retrospective methods, it is capable of estimating the  $z$  term, which helped it to perform well consistently on every data set. In terms of correction quality, it surpassed all tested methods including the gold standard, and it also substantially reduced errors in cell counting, cell intensity and cell area measurements (**Supplementary Fig. 1**). CIDRE is available as open-source software in Matlab (<http://www.github.com/smithk/cidre/>) and as an ImageJ plug-in (**Supplementary Software**).

Although CIDRE is useful for many applications, it is not suitable in certain conditions. The key assumption is violated if the images are highly correlated. In time-lapse images, for example, this may cause artifacts in the correction, although combining images from different sites can help reduce this danger. Like other retrospective methods, CIDRE performs best with many images containing ample intensity information. With 1,000 images or more, CIDRE was substantially better than the gold standard on average (**Fig. 2** and **Supplementary Fig. 2**). Fewer images or sparse intensity information reduced correction quality. We found that for most applications, ten images were sufficient to ensure an improvement in illumination quality, whereas approximately 100 images were necessary to match the gold standard (**Supplementary Fig. 3**).

Uneven illumination is a common but misunderstood phenomenon, and results reported by many researchers have undoubtedly been affected by it. Proper illumination-correction procedures should be followed whenever acquiring images for quantitative microscopy, and CIDRE is a simple, freely available tool that can help ensure the quality of such measurements.

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

*Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).*

## ACKNOWLEDGMENTS

We thank R. Dechant, T. Schwarz, A. Kaufmann and J. Kusch for data collection; L. Lenherr Smith for help with the manuscript; and our colleagues who completed the survey. This work was funded by SystemsX.ch, the Swiss national initiative for Systems Biology, through the SyBIT project. P.H. acknowledges support from the Finnish TEKES FiDiPro and the Hungarian National Brain Research Programme (MTA-SE-NAP B-BIOMAG).

## AUTHOR CONTRIBUTIONS

P.H., F.P., A.B. and G.C. initiated the project. K.S. and P.H. designed the correction method. K.S. and Y.L. designed the optimization. K.S. and C.B. implemented the software. F.P. and K.S. collected the image data. K.S. performed the analysis and wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

**Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.**

1. Goldman, D.B. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 2276–2288 (2010). [Medline](#) [CrossRef](#)
2. Carpenter, A.E. *et al. Genome Biol.* **7**, R100 (2006). [Medline](#) [CrossRef](#)
3. Young, I.T. *Curr. Protoc. Cytom.* **14**, 2.11 (2001).
4. Model, M. *Curr. Protoc. Cytom.* **68**, 10.14 (2014).
5. Ljosa, V. & Carpenter, A.E. *PLoS Comput. Biol.* **5**, e1000603 (2009). [CrossRef](#)
6. Shariff, A., Kangas, J., Coelho, L.P., Quinn, S. & Murphy, R.F. *J. Biomol. Screen.* **15**, 726–734 (2010). [Medline](#) [CrossRef](#)

**Figure 1** | Summary of illumination correction using CIDRE. **(a)** The observed image is corrupted by misaligned optics, vignetting, dust, etc., and this corruption is modeled as a linear intensity gain function  $v$ . The observed image also contains contributions from camera offset and thermal sources, modeled as an additive term  $z$ . **(b)** To recover the true image, we consider the local distributions of observed intensities (red and blue) collected from many images. Each is related to the true underlying distribution of the specimen (gray, upper plot) by a linear transform parameterized by  $v$  and  $b$ , which correspond to slope and  $y$  intercept, respectively, in a quantile-quantile plot between the true distribution and the local distribution. **(c)** The plot shows an estimation of the parameters in **b** simultaneously for all locations by minimizing a regularized energy function composed of several terms. **(d–f)** Terms composing the energy function in **c**. The plot in **d** shows a robust regression term that ensures parameter values fit the data. The plot in **e** shows a regularization term that reduces noise and guarantees smoothness of the correction surface, forcing neighboring distributions (green and blue) to agree on similar parameter values. The plot in **f** shows how  $z$ , a nearly uniform-intensity surface representing the intensity recorded when no light is present, is estimated by finding the common point where all regression lines intersect. **(g)** When we apply the reverse transform using estimates of  $v$  and  $z$ , local-intensity distributions (red and blue) take the shape of the true distribution (gray, upper and lower plots), and the uncorrupted images can be recovered.

**Figure 2** | Comparison of 13 illumination correction methods, averaged over 12 image collections. Scores are the mean absolute differences between pairs of overlapping test images after correction, computed pixelwise and normalized by the disagreement between uncorrected image pairs. A score of 1 implies equivalent disagreement to the uncorrected image pairs, indicated by dashed lines. A score of 0 implies a perfect correction, though it is not achievable in practice. <sup>†</sup>Prospective methods. <sup>\*</sup>Single-image methods. <sup>‡</sup>Gold-standard methods. Methods and data sets are described in the Online Methods.

## ONLINE METHODS

**Software.** Our software, CIDRE, is available as an ImageJ plug-in and as a Matlab script. The open-source code for CIDRE is available as **Supplementary Software** and for download at <http://www.github.com/smithk/cidre/>. Compiled versions of the code and updates are available at <http://www.scopem.ethz.ch/research/software/cidre.html>.

**Microscopy data sets.** We collected 11 data sets designed to represent a variety of microscopy setups commonly used for research to evaluate illumination-correction methods (**Supplementary Table 1**). Misaligned optics, dust, scratches, vignetting, and nonuniform light sources contribute to uneven illumination in these data sets. The most common source of uneven

illumination is vignetting, a radial intensity falloff from the center of the image<sup>7</sup> caused by blockages of extreme light paths, varying angular sensitivity of the sensor, and natural geometric falloff<sup>1</sup>. Images were acquired with eight different microscopes and one conventional SLR camera. We varied the microscope itself, sample preparation, staining, mounting, objectives, light sources, type of sensor, image resolution, and bit depth. In addition, we generated a set of synthetic images to form a twelfth data set allowing us to directly measure correction quality. Each data set contains several sets of images: approximately 100 images of a reference material (for example, fluorescent slide), 10–100 ‘empty’ images (for example, the culture medium or the glass slide), 10–100 dark-frame images where no light is incident on the sensor (acquired with the light source switched off or otherwise blocked), 10–100 images in a series with increasing exposure times, 100–200 pairs of images used to test the correction models, 1,000–3,000 images used to train the correction models, and 1,000–3,000 segmentations of the training images obtained with CellProfiler<sup>2</sup>. Below we give a brief description of each data set (**Supplementary Note 2**).

**HCS-DNA:** this data set contains propidium iodide (P3566 Invitrogen)-stained ATCC HeLa cells from a high-content screen. The compound stains DNA and RNA, visualizing the nuclei and cytoplasm as well as clumps of Semliki Forest virus. Images were collected on the RFP channel of a fluorescence wide-field microscope using 384-well plates with a plastic-air interface. The microscope was a Molecular Devices ImageXpress Micro with a 20×/0.75-NA objective with a LED-liquid light guide (Lumencor Spectra X). A CCD sensor (Photometrics CoolSNAP HQ) acquired  $1,392 \times 1,040$  12-bit images. The mean signal intensity in the images was low (219). Reference images were acquired at various locations on a plastic fluorescent reference slide and empty wells containing fluorescent medium. The intensity attenuation measured as the difference between the 99th percentile and 1st percentile intensities computed from the mean of the reference images was 42.6%. The confluency, measured as the percentage of the image area occupied by cells, was 21.5%.

**HCS-NUC:** this data set contains DAPI-stained MRC5 human fetal lung fibroblasts (SV40 immortalized) from a high-content screen. Images were collected using a second microscope with a similar setup to that of HCS-DNA but with a DAPI-adapted filter set. The mean signal intensity was moderate (549), intensity attenuation was 18.7%, and cell confluency was low (2.7%).

**HCS-ACTIN:** this data set contains actin-stained HeLa cells from a high-content screen. Images were collected using the same microscope and setup as for HCS-NUC, except that a different filter set and light source were used (Lambda LS arc lamp, Sutter Instruments, with a 300-W xenon bulb, PerkinElmer). The mean signal intensity was bright (1,776), intensity attenuation was 23.3%, and confluency was high (88.2%).

**MICROFLUID:** this data set contains yeast cells in a synthetic medium containing glucose and dextran coupled with Alexa Fluor 680, used to study regulation of the PKA pathway through V-ATPase. A microfluidic setup was used (CellAsic Microfluidic plate Y04C), and images were acquired using a Nikon Ti-Eclipse microscope with a 60×/1.4-NA objective and LED-liquid light guide (CoolLED pE-2). A high-resolution sCMOS sensor (Hamamatsu Orca Flash 4.0) was used to acquire  $2,048 \times 2,048$  16-bit images. Reference images were acquired from various locations on a plastic fluorescent reference slide and empty wells containing fluorescent medium. The mean signal intensity was bright (1,086), intensity attenuation was very strong 50.4%, and cell confluency was very low (0.6%).

**FLUO-ACTIN:** this data set contains muntjac skin fibroblast cells on a glass slide (FluoCells prepared slide F36925). The prominent filamentous actin in these cells was labeled with Alexa Fluor 488. Images were collected using a fluorescence wide-field microscope with a glass-oil interface. The microscope model was a Zeiss 200M with a 63×/1.4-NA objective and a HBO-direct light source with a 100-W Mercury arc lamp. A CCD sensor (Hamamatsu C8484) acquired  $1,344 \times 1,024$  12-bit images. Reference images were acquired at various locations on a plastic fluorescent reference slide. The mean signal intensity was low (213), intensity attenuation was weak (7.5%) and cell confluency was moderate (32.5%).

**FLUO-MITO:** this data set contains bovine pulmonary artery endothelial (BPAE) cells on a glass slide (FluoCells prepared slide F36924). Mitochondria were labeled with red fluorescent MitoTracker Red CMXRos. Images were collected using a fluorescence wide-field microscope with a glass-air interface. The microscope model was an Olympus IX81 with a 20×/0.75-NA objective and a Lambda LS arc lamp (Sutter Instruments) with a 175-W xenon bulb (PerkinElmer). The CCD sensor (Hamamatsu Orca ER) acquired  $1,344 \times 1,024$  12-bit images. Reference images were acquired at various locations on a plastic fluorescent reference slide. The mean signal intensity was low (338), intensity attenuation was strong (40.3%), and cell confluency was low (13%).

**FLUO-EMCCD:** this data set contains bovine pulmonary artery endothelial (BPAE) cells on a glass slide (FluoCells prepared slide F14781). Anti-bovine alpha-tubulin mouse monoclonal in conjunction with goat anti-mouse antibodies label the microtubules. Images were collected using a fluorescence wide-field microscope with a glass-air interface. The microscope model was a Leica DMI6000B with a 40×/0.7-NA objective and a 100-W metal-halide lamp. An EMCCD sensor (Andor iXon 885) acquired  $1,004 \times 1,002$  16-bit images with the EM gain set at a medium setting (3). Reference images were acquired at various locations on a plastic fluorescent reference slide. The mean signal intensity was bright (2,155), intensity attenuation was 38%, and cell confluency was moderate (34.3%).

**HIST-CONFOCAL:** this data set contains a 16- $\mu$ m cryostat section of mouse kidney on a glass slide stained with Alexa Fluor 488 to label elements of the glomeruli and convoluted tubules (FluoCells prepared slide F24630). Images were collected using a fluorescence confocal microscope with a glass-air interface. The microscope model was a Zeiss LSM 510 with a 10×/0.3-NA objective and a 30-mW argon 488-nm laser with an optical fiber light guide. A Hamamatsu photomultiplier tube (PMT) sensor acquired  $512 \times 512$  12-bit images. Reference images were acquired at various locations on a plastic fluorescent reference slide. The mean signal intensity was medium (614), intensity attenuation was strong (43.9%), and confluency was high (54%).

**HIST-BRIGHT:** this data set contains a sagittal section of mouse brain on a glass slide labeled with hematoxylin and eosin (HE) stain. Bright-field images were collected using a wide-field microscope with a glass-oil interface. The microscope model was an Olympus IX81 with a 40×/1.3-NA objective and Lambda LS arc lamp (Sutter Instruments) with a 175-W xenon bulb (PerkinElmer). A CCD sensor (Hamamatsu Orca ER) acquired  $1,344 \times 1,024$  12-bit images. Reference images were acquired using empty locations on the glass slide. The mean signal intensity was bright (614), and intensity attenuation was 32%.

**PHASE:** this data set contains native epithelial cells from oral mucosa on a glass slide. Phase-contrast images were collected using a wide-field microscope with a glass-air interface.



The microscope model was an Olympus IX81 with a 10×/0.3-NA objective and 100-W halogen lamp. A CCD sensor (Hamamatsu Orca ER) acquired  $1,344 \times 1,024$  12-bit images. Reference images were acquired using empty locations on the glass slide. The mean signal intensity was moderate (923), intensity attenuation was 12.4%, and cell confluency was low (7.1%).

**PHOTOGRAPHY:** this data set contains conventional photographs of projections of paintings and drawings from the National Gallery of Art. A consumer digital SLR camera was (Nikon D90) was used with an 18- to 105-mm zoom lens (set at 50 mm). Reference images were acquired using empty sheets of white paper. A CMOS sensor (Nikon DX format) acquired  $1,072 \times 712$  8-bit images. The mean signal intensity was bright (150), and intensity attenuation was 17%.

**SYNTHETIC:** this data set contains synthetically generated images of HeLa cells. Uncorrupted  $1,392 \times 1,040$  12-bit images were synthetically generated and artificially subjected to illumination distortions using a physically plausible model (**Supplementary Note 6**). Reference images were synthetically generated in a similar manner. Because the corruption process is known and the image pairs are perfectly aligned, this data set allows us to directly measure differences between corrections and uncorrupted images.

**Baseline illumination-correction methods.** We compared our method to 12 commonly used prospective and retrospective approaches for illumination correction (**Supplementary Note 3**). Each method is summarized below, starting with prospective correction methods, which require special calibration images to be acquired along with the data.

**CALIB-ZERO** models the illumination gain  $v$  as the average of a set of images of a plastic fluorescent reference slide<sup>8</sup> at various locations. The zero-light term  $z$  is modeled by averaging a set of images taken at various locations with the shutter closed, the light source turned off or otherwise blocked. Image correction follows equation (1), reversing the image formation process. This approach is considered a gold standard because it empirically models both correction terms. However, it has several drawbacks (**Supplementary Note 3**).

**CALIB** models the illumination gain  $v$  as above but ignores the zero-light term<sup>8</sup>. Image correction is done by normalizing the image by the gain  $I(x) = I_0(x)/v(x)$ . This is an incomplete model, but it can give reasonable results when the zero-light term is small relative to the calibration intensity.

**EMPTY-ZERO** models the illumination gain  $v$  as the average of empty' images taken at various locations<sup>3</sup>. In our experiments, this includes either the culture medium without any cells or the glass slide. The zero-light term  $z$  is modeled by averaging a set of images taken with the shutter closed. Image correction reverses the image formation process (equation (1)). This approach also serves as a gold standard because it empirically models the correction terms, but it suffers from the same drawbacks as CALIB-ZERO (**Supplementary Note 3**). This method is appropriate for bright-field images or fluorescence imaging when the medium fluoresces.

**EMPTY** models the illumination gain  $v$  as above but ignores the zero-light term<sup>8</sup>. Image correction is done by normalizing the image by the gain  $I(x) = I_0(x)/v(x)$ . Although flawed, this approach can give reasonable results when  $z$  is small.

**SIGAL** uses a specially acquired set of images of a homogeneous medium captured with increasing exposure times to build the correction model<sup>9</sup>. Estimates of  $v$  and  $z$  are obtained by performing a least-squares fitting on data from the exposure series, giving additional weight to

data from zero exposure time to anchor the zero-light level. Image correction reverses the image formation process (equation (1)). Because the exposure series often contains only a few data points (7 in the original paper<sup>9</sup>, 13–100 in our data) and are not regularized, estimates of  $v$  and  $z$  tend to be noisy and non-smooth.

Next, we describe four purely computational retrospective methods that build the correction model from more than one image.

CAN builds a model of the illumination gain by sorting image data by location<sup>10</sup>. The top 10th-percentile data are used to fit a polynomial surface in the log-intensity space to estimate  $v$ . Images are corrected according to  $I(x) = I_0(x)/v(x)$ . This approach ignores the zero-light term, but it can give reasonable results when that term is small relative to the signal.

LJOSA forms the illumination gain estimate  $v$  by fitting a polynomial or spline surface to an average of the images to be corrected<sup>5</sup>. The fitted surface is used for correction according to  $I(x) = I_0(x)/v(x)$ . This approach ignores the zero-light term, but it can give reasonable results when that term is small relative to the signal.

SCHWARZFISCHER is a method designed to correct time-lapse fluorescence images, though it can be used for static images if the background contains illumination information<sup>11</sup>. Each image is broken into overlapping tiles, and each tile is clustered as either background or signal on the basis of local statistics. Next, background models are constructed for each image using natural-neighbor interpolation to smooth and fill in missing values. A foreground model is estimated using least-squares fitting at every location on the background data. The images are corrected with a formula similar to (1). This method can compensate for photobleaching as well as uneven illumination, but it becomes unreliable for high-confluency images because it depends on a large background area (which is sparse when many cells are present). Comparing intensities between images corrected with this method is problematic because the correction model is adapted to each image.

SHARIFF is the default method<sup>6</sup> used in the software package CellProfiler<sup>2</sup>, which supports several options for illumination correction. The intensity gain is estimated using the mean image intensity computed at every location. It is smoothed with a median filter and rescaled to an appropriate range. Images are corrected according to  $I(x) = I_0(x)/v(x)$ .

Finally, we describe three single-image retrospective correction methods. These methods are applied independently to every image and do not require any calibration images. Although they are useful for improving the appearance of the image, these methods are inappropriate if intensity measurements will be made from the image because they nonuniformly alter the signal.

ROLLINGBALL or the top-hat transform applies a shape filter that subtracts a geometric opening of the image from the image<sup>12</sup>. The ‘ball’ in this case refers to the shape of the filter kernel, which is sized to be larger than the diameter of the largest expected object. For each data set, we tested several kernel sizes and chose the one with the best performance.

SMOOTHING also relies on subtracting a filtered version of the image from the original<sup>13</sup>. In this case, a Gaussian smoothing kernel is used, which is sized to be larger than the diameter of the largest expected object.

BABALOUKAS selects locations from the background of the image, fits a polynomial to those points, and subtracts the fitted surface from the original image<sup>14</sup>. In the original paper,

background points were selected manually. We automatically selected background points using the clustering technique proposed by Schwarzfischer *et al.*<sup>11</sup>.

**Evaluation protocol.** To assess the quality of correction (**Fig. 2** and **Supplementary Figs. 2 and 3**), we defined an evaluation procedure, error measure, and benchmark score. Measuring correction quality is difficult because we cannot directly compare against the uncorrected images. Some authors use the coefficient of variation (CV) of corrected reference images<sup>15</sup>, but differences between the reference images and actual data cast doubt on the validity of this measure. Our solution was to collect hundreds of pairs of overlapping images for each data set, precisely align them, and report the mean of absolute differences for each pair of corresponding pixels in the overlapping regions (**Supplementary Note 5**). We collected 100–200 pairs of test images for every data set. The various correction methods were applied to each data set.

Prospective methods used sets of reference images to train the model, whereas multi-image retrospective methods were trained using a collection of 1,000–3,000 images acquired under similar conditions as the test set. Single-image retrospective methods operate directly on the test images. Each pair of test images was aligned using a subpixel registration technique<sup>16</sup>. The overlapping regions from each test pair were used in the evaluation (**Supplementary Note 5**). Whenever possible, we used a four-quadrant solution wherein quadrant IV of the first image overlapped quadrant II of the second image. Our reasoning is that the most dramatic intensity difference often occurs between a corner of the image and the center, and this format allowed us to directly compare the corner of one image to the center of the other. However, our ability to automate the data collection process was limited by the microscope software and size of the specimen, so this layout was not always possible. In two cases (FLUO-MITO and PHASE), we were forced to overlap the left half of the first image with the right half of the second image. In the case of the SYNTH data set, each pair of test images is perfectly aligned, and so the entire image overlaps. For a given test pair, the overlap regions are two separate acquisitions of the same field of view, denoted IM1 and IM2. The principal sources of disagreement between IM1 and IM2 are uneven illumination, noise, alignment errors, and photobleaching. The error is the mean of absolute differences between all aligned pixel pairs in the overlapping regions. More precisely

$$e = \frac{1}{N} \sum_n \sum_l |I_1^n(l) - I_2^n(l)|$$

where  $e$  is the error,  $n$  indexes the image pair,  $l$  indexes locations in the overlapping region, and  $N$  is the total pixel pairs across all images pairs. To correct for photobleaching, and because the outputs of the correction methods have different dynamic ranges (some methods are purely multiplicative, some purely subtractive, etc.), we standardized the intensity distributions of the images, thus ensuring comparability across correction methods. The median and s.d. of the set of test image sets were normalized to those of the uncorrected IM1 image set. Finally, the mean error for all image pairs is normalized by the benchmark error, that of the uncorrected image pairs, with

$$s_{\text{method}} = \frac{\bar{e}_{\text{method}}}{\bar{e}_{\text{UNCORRECTED}}}$$

This yields a score  $s$  in the interval  $[0, \infty)$ , where 0 indicates perfect correction (IM1 and IM2 are the same), 1 indicates disagreement equivalent to the uncorrected images, and  $>1$

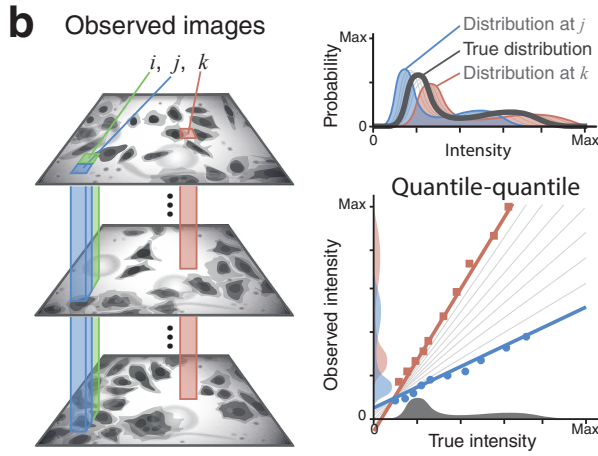
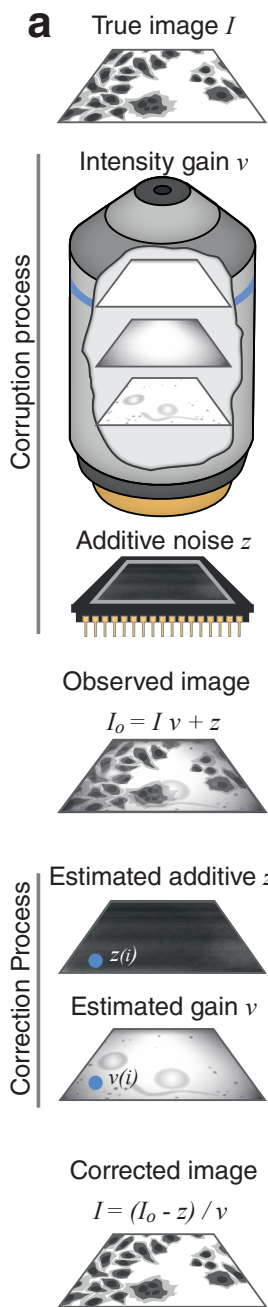
indicates greater disagreement than the uncorrected images (i.e., the correction process increased disagreement between image pairs).

**Assumptions.** Our correction method makes the following assumption: the probability of the specimen appearing at location  $x$  in an image is uniform for all  $x$ , and there is no correlation between the images. In other words, the content of the image can appear anywhere in the image with equal probability. For example, a cultured cell or a piece of tissue are just as likely to appear in the corner of the image as in the center. A second assumption is based on empirical observations: the zero-light noise  $z(x)$  should be approximately uniform for all  $x$  (flat across the image). This is based on evidence from measurements of the zero-light noise from the data we collected and reliable sensor manufacturing. To verify this assumption we captured and averaged dark frames from each data set. We found that the s.d. of such images was less than 0.1% of the dynamic range, and often less than 0.02%.

**Illumination-correction formulation.** Our method for illumination correction, CIDRE (corrected intensity distributions using regularized energy minimization), operates on a set of observed images corrupted by illumination distortion that were acquired under similar conditions. CIDRE recovers the uncorrupted images by first estimating the unknown parameters  $v$  and  $z$  through an energy-minimization technique and then applying equation (1) to recover the uncorrupted images (**Supplementary Note 4**). It is assumed that objects may appear anywhere in the image with equal probability. If this assumption is valid, then the distribution of intensities at a single location taken from infinitely many images is related to an underlying distribution common to every location by a linear transform modeling the corruption process. Local-intensity distributions from a finite set of observed images are simply linear transforms of a sampling of the underlying distribution parameterized by  $v$ ,  $b$  and  $z$ . CIDRE estimates these parameters simultaneously for all locations by using a quasi-Newton method<sup>17</sup> to minimize a regularized energy function comprising a robust regression term, a smoothness term, and an offset term. The robust regression term ensures  $v$  and  $b$  fit the data (**Fig. 1d**). The smoothness term enforces a smooth correction surface and reduces noise by encouraging neighboring distributions to agree on similar values for  $v$  (**Fig. 1e**). The offset term estimates  $z$  by finding the common point where all regression lines intersect (**Fig. 1f**). The results of the optimization are robust estimates for  $v$  and  $z$ , which allow us to apply equation (1) and recover the uncorrupted images.

7. Ray, S.F. *Applied Photographic Optics* 3rd edn. (Focal, 2002).
8. Varga, V.S. *et al. Cytometry A* **60**, 53–62 (2004). [Medline](#) [CrossRef](#)
9. Sigal, A. *et al. Nat. Methods* **3**, 525–531 (2006). [Medline](#) [CrossRef](#)
10. Can, A. *et al. in Proc. IEEE Int. Symp. Biomed. Imaging* 288–291 (IEEE, 2008).
11. Schwarzfischer, M. *et al. in Micro. Image Anal. Appl. Biol.* (MIAAB, Heidelberg, Germany, 2011).
12. Sternberg, S.R. *Computer* **16**, 22–34 (1983). [CrossRef](#)
13. Leong, F.J., Brady, M. & McGee, J.O. *J. Clin. Pathol.* **56**, 619–621 (2003). [Medline](#) [CrossRef](#)
14. Babaloukas, G., Tentolouris, N., Liatis, S., Sklavounou, A. & Perrea, D. *J. Microsc.* **244**, 320–324 (2011). [Medline](#) [CrossRef](#)
15. Olsen, D. *et al. Remote Sens.* **2**, 464–477 (2010). [CrossRef](#)

16. Guizar-Sicairos, M., Thurman, S.T. & Fienup, J.R. *Opt. Lett.* **33**, 156–158 (2008).  
[Medline](#) [CrossRef](#)
17. Liu, D.C. & Nocedal, J. *Math. Program.* **45**, 503–528 (1989). [CrossRef](#)



**c** Energy minimization simultaneously estimates  $v$  and  $z$  at all locations

$$E = E_f + E_r + E_z$$

