Background intensity correction for terabyte-sized time-lapse images

J. CHALFOUN*, M. MAJURSKI*, K. BHADRIRAJU†, S. LUND*, P. BAJCSY* & M. BRADY*

**Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

†*Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

Summary

Several computational challenges associated with large-scale background image correction of terabyte-sized fluorescent images are discussed and analysed in this paper. Dark current, flat-field and background correction models are applied over a mosaic of hundreds of spatially overlapping fields of view (FOVs) taken over the course of several days, during which the background diminishes as cell colonies grow. The motivation of our work comes from the need to quantify the dynamics of OCT-4 gene expression via a fluorescent reporter in human stem cell colonies. Our approach to background correction is formulated as an optimization problem over two image partitioning schemes and four analytical correction models. The optimization objective function is evaluated in terms of (1) the minimum root mean square (RMS) error remaining after image correction, (2) the maximum signal-to-noise ratio (SNR) reached after downsampling and (3) the minimum execution time. Based on the analyses with measured dark current noise and flat-field images, the most optimal GFP background correction is obtained by using a data partition based on forming a set of submosaic images with a polynomial surface background model. The resulting image after correction is characterized by an RMS of about 8, and an SNR value of a 4×4 downsampling above 5 by Rose criterion. The new technique generates an image with half RMS value and double SNR value when compared to an approach that assumes constant background throughout the mosaic. We show that the background noise in terabyte-sized fluorescent image mosaics can be corrected computationally with the optimized triplet (data partition, model, SNR driven downsampling) such that the total RMS value from background noise does not exceed the magnitude of the measured dark current noise. In this case, the dark current noise serves as a benchmark for the lowest noise level that an imaging system can achieve. In comparison to previous work,

the past fluorescent image background correction methods have been designed for single FOV and have not been applied to terabyte-sized images with large mosaic FOVs, low SNR and diminishing access to background information over time as cell colonies span entirely multiple FOVs. The code is available as open-source from the following link https://isg.nist.gov/.

Background

Pluripotent stem cells have great potential as a source of cells for regenerative therapies. However, many aspects of controlling pluripotent stem cells behaviour are still not well understood (Saha & Jaenisch, 2009). The motivation of our work comes from the need to quantify the dynamics of OCT-4 gene expression in human stem cell colonies, because OCT-4 is a critical gene in the regulation of pluripotency, or the ability of stem cells to differentiate into all somatic cell types (VanDenBerg *et al.*, 2010). OCT-4 gene expression in cells is reported by a green fluorescent protein (GFP) reporter inserted in the regulatory region of the OCT-4 gene (Zwaka & Thomson, 2003). Specifically, we are interested in understanding how colony-level GFP intensity is related to population-level cell behaviour, how normal regulation of stem cell gene expression occurs, and how to develop and assess human pluripotent stem cells culture quality parameters.

Time-lapse epifluorescence microscopy using fluorescent protein reporters provides an opportunity for imaging and analysing the dynamics of gene expression and morphological changes in live human pluripotent stem cells cultures. Imaging at high spatial and temporal resolutions generates terabyte-sized image sets spanning hundreds of FOVs through time (Fig. 1). There are several technical challenges to overcome before quantitative biological information can be obtained from these big data sets. Images of live cells, such as pluripotent stem cells, must be acquired with low-power illumination to minimize biological artefacts from light-induced damage to cells. Light intensity was empirically set at the lowest value, at which we could still discern image features in each

Correspondence to: Peter Bajcsy, Information Technology Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, U.S.A. Tel: 3019752958; fax: 301-975-6097; e-mail: peter.bajcsy@nist.gov
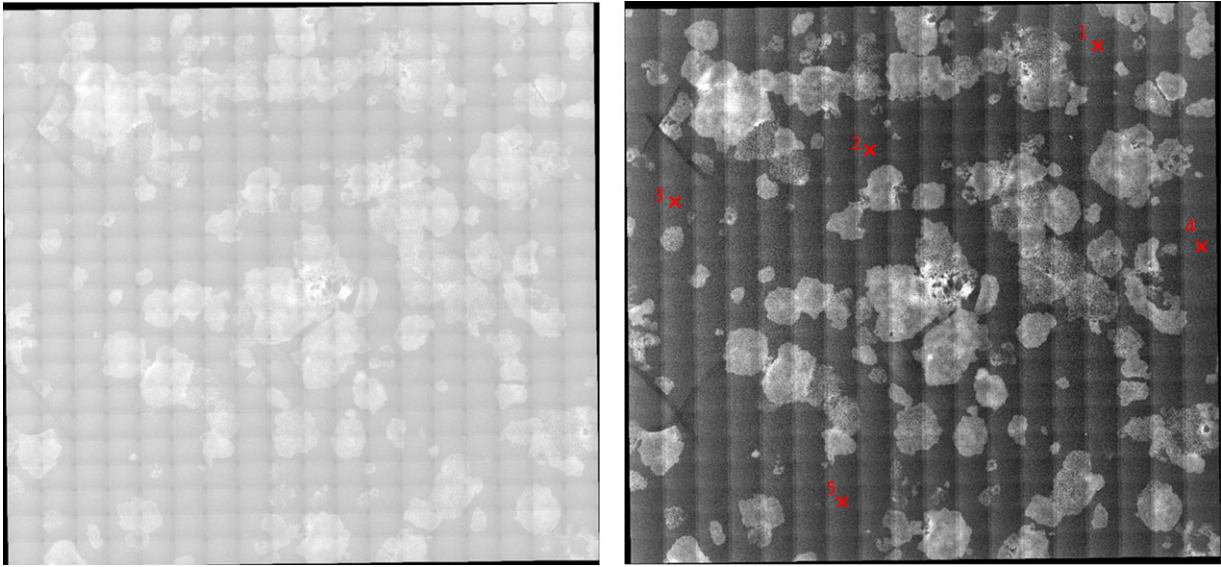
Fig. 1. Uncorrected and flat-field corrected example image: Uncorrected mosaic of green fluorescent protein (GFP) image channel that was stitched from 22×18 FOVs acquired on the fourth day of a stem cell colony imaging experiment (left). The image on the right is corrected for uneven illumination and dark current but not corrected for background. The five numbered red 'x' mark FOVs with nothing but background pixels in them.
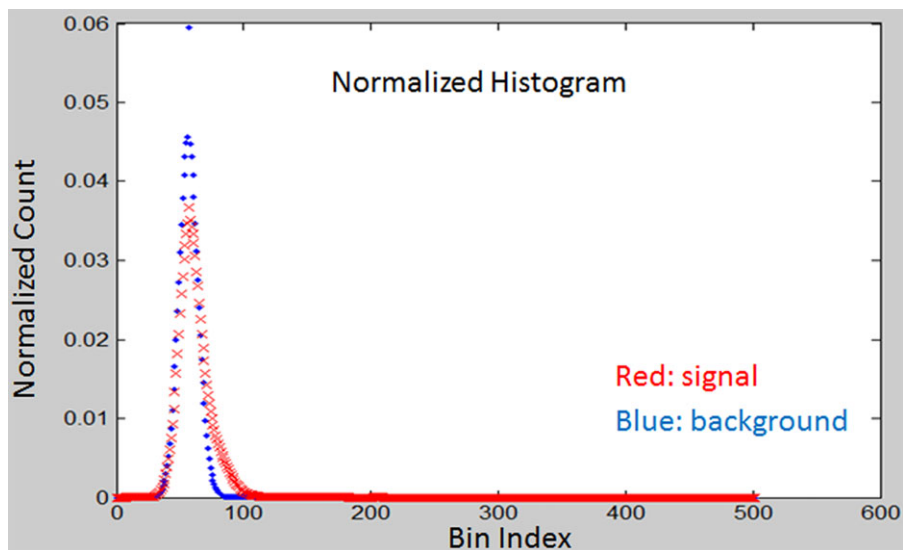

Fig. 2. Normalized histogram of foreground and background intensities: Foreground and background normalized histograms of one image with 500 bins. The two curves illustrate low SNR = 1.4 in the (1×1) binned GFP image.

respective colour channel and could perform image analysis. The acquisition requirement of minimally perturbing cells leads to a low signal-to-noise ratio (SNR) of the fluorescent signal (Fig. 2) and its sensitivity to correcting for dark current, flat-field and background media sources of noise. We address the overall problem of fluorescent image correction with the focus on background correction in order to minimize the remaining errors in the corrected background and maximize the SNR.

The background correction poses challenges due to the complex interactions of cells, media, fluorescent biomarker and imaging light, and also due to the computational demands of processing images spanning very large fields of view of growing cell colonies. In general, the foreground intensity signal over cell colony image regions includes not only contributions from cells but also non-specific autofluorescence from the cell culture media, culture dish and any extracellular matrix protein coatings (Fig. 3). The fluorescent components may vary
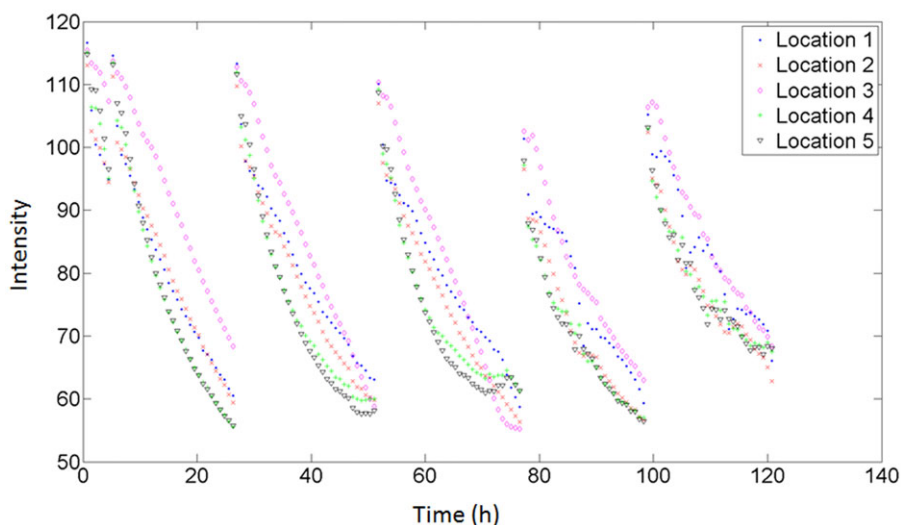
Fig. 3. Spatiotemporal graphs of average intensity per FOV: These results are displayed for the five numbered locations shown in Figure 1 through 5 days of acquisition. A frame is acquired every 45 min and a daily media change for the culture is performed every 24 h from the beginning of the experiment. The curves illustrate global intensity variations across an entire image mosaic and over time.

spatially and temporally because of spatial distribution of fluorescent molecules, different amounts of light interacting with these molecules, and photo-bleaching of molecules over time. It is impossible to capture these spatiotemporal variations of fluorescent signal in the background from a single FOV since stem cell colonies grow to cover very large spatial areas. Furthermore, the background correction is complicated by the fact that individual colonies can completely occupy areas spanning several FOVs, creating FOVs without any background pixels (Fig. 1). Unlike single cell imaging where background areas around cells provide accurate estimates of background intensity, pluripotent stem cells grow as colonies of cells (islands) that merge with neighbouring colonies over time as the culture progresses, and background areas in a colony culture become sparse at later times. In summary, background correction from live stem cell colony microscopy experiments has to overcome estimation challenges including: (1) finding accurate spatial and temporal models, (2) deriving parameters of the models from a diminishing number of background pixels over time as colonies grow, (3) maximizing SNR due to low-light illumination and (4) scaling computations of background models over a very large mosaic of small FOVs, and hence millions of pixel values per time frame.

The problem of correcting single FOV images has been approached by imaging reference materials (Model & Burkhardt, 2001), designing data-driven models of optical and digital artefacts without recorded images of reference materials (Leong, 2003), or combining both recorded images of reference materials and data-driven models (Piccinini *et al.*, 2013). The past work spans commercial panoramic cameras (Goldman, 2010), microscope cameras (Waters & Swedlow, 2007; Wu *et al.*, 2008), and various custom cameras (Kim

& Pollefeys, 2008; Galego, 2011). Many techniques for estimating vignetting (radial fall-off), exposure and white balance variations, and sample radiances are common across multiple camera types including the microscope cameras with specific challenges summarized by Waters & Swedlow (2007). Some background correction models are approximated by the mean grey level of the areas between cells (Model & Burkhardt, 2001). A more accurate background subtraction can be achieved by extracting average grey level from predefined surrounding pixels of a cell or a colony (Chalfoun *et al.*, 2013). This cell/colony specific background value compensates for spatial variations better than a single mean grey level especially for data sets with large FOVs and acquired over a significant duration of time.

None of the previous work deals with background measurement or modelling over large image mosaics and in the presence of cells/colonies that cover multiple FOVs. The background is usually considered constant or is neglected. In more advanced treatment, background is approximated by averaging closest background pixels to the cell areas or estimated by a fitting function on the entire FOV pixels. These approaches do not work well on large spatial mosaics due to computational complexity (execution time and machine memory usage) and the need to model the background across large numbers of FOVs.

Our approach to background correction is formulated as an optimization problem over two data partitioning schemes (full mosaic and submosaic based) and their corresponding correction models represented by four analytical functions. The four functions include polynomial surface, cubic spline interpolation, linear interpolation and nearest neighbour. The optimization objective function is evaluated in terms of (1) the

minimum RMS error remaining after correction, (2) the maximum SNR reached after downsampling and (3) the minimum execution time and low memory usage. In terms of SNR maximization, a successful background correction produces an image where the background follows a standard white noise reduction by a factor of n when the image is binned by $n \times n$. The method that has the maximum SNR after downsampling with the same binning factor of $n \times n$ will be better suited for this type of correction.

Based on the analyses with measured dark current noise and flat-field images, we show that the most optimal GFP background correction is obtained by using the submosaic-based data partition and a polynomial surface background model. It is critical in low-SNR images that the remaining error after GFP correction is minimized in order to derive GFP signal-based biological insights with high statistical confidence. We show that the background noise contributing to the total RMS value can be corrected computationally with the developed method down to the magnitude of the measured dark current noise. In this case, the dark current noise serves as a benchmark for the lowest noise level that an imaging system can achieve. In comparison to previous work, the past GFP correction methods have been designed for a single FOV and have not been applied to terabyte-sized images with low SNR, colonies spanning multiple FOVs and missing background information. Thus, the novelty of our work is in the design of background correction of large-sized GFP images such that the amount of remaining errors after correction is minimized and the SNR is maximized whereas the spatial resolution is minimally sacrificed.

This paper describes a search for an optimal background correction triplet (data partition, model, SNR driven downsampling) over terabyte-sized fluorescent image mosaics such that the noise from background after correction does not exceed the magnitude of measured dark current noise.

The outline of the paper is as follows. Section 2 presents all the steps for GFP image correction including the data partitioning schemes and the surface fit functions used to perform the background correction. Section 3 documents all evaluation results to find the optimal data partitioning and fit function with low SNR in the image. Sections 4 and 5 summarize and discuss our results.

Methods

We proceed under the assumption that the GFP foreground and background masks are available over all time frames. In our work, these masks were created from phase contrast microscopy images by thresholding gradient images obtained via Sobel operator and postprocessed using morphological operations that include deleting objects smaller than 5000 pixels. The threshold value is selected by using an empirical gradient threshold (EGT) technique as documented in (Chalfoun et al.,

2014). Two experts performed manual segmentation on 16 images with a frequency of 1 image every 10 frames to cover the entire time-sequence. The comparison between manual and automated segmentations was measured using the Dice metric (Dice, 1945). An average value of 0.9944 ± 0.0043 for Dice is obtained between manual and automated segmentation of the phase contrast images with a clear separation between foreground (colonies) and background. Results of a typical segmentation from this automated process are shown in Figure 4. Both phase contrast and GFP images are registered by virtue of the microscope design.

Image acquisition

Our dataset includes three experimental replicates following the growth in culture of the H9 human embryonic stem cell line over 5 days on a microscope stage equipped with a controlled environment incubation chamber (Kairos Instruments LLC, Pittsburgh, PA, USA). The stem cell line was engineered to stably produce green fluorescent protein (GFP) under the influence of the native OCT-4 promoter, using a published homologous recombination plasmid construct developed by the James Thomson lab (Zwaka & Thomson, 2003), and obtained from Addgene (Addgene, Cambridge, MA, USA).

Cells were cultured in phenol red-free DMEM/F12 supplemented with sodium selenium (14 ng mL^{-1}), insulin (19.4 μgmL^{-1}), sodium bicarbonate (543 μgmL^{-1}), transferrin (10.7 μgmL^{-1}) (all from Life Technologies, Carlsbad, CA, USA), L-ascorbic acid-2-phosphate magnesium (64 μgmL^{-1}; Sigma-Aldrich, St. Louis, MO, USA), basic fibroblast growth factor (100 ng/mL), and transforming growth factor-β1 (2 ng mL^{-1}) (both from R&D systems, Minneapolis, MN, USA) (Chen et al., 2011).

Imaging was done on a Zeiss 200M microscope using a Zeiss 10$\times$ EX Plan-NeoFluar ph1 0.3NA objective (Carl Zeiss Microscopy, LLC, Thornwood, NY, USA) both in phase contrast and epifluorescence modes, with mosaics acquired every 45 min. Images were captured on a Coolsnap fx camera (Photometrics, Tucson, AZ, USA) with 12-bit digitization and no binning. Phase contrast images were captured using a 150 ms exposure with a 100 W tungsten–halogen light source set at 2.3 V. Green fluorescence images were captured using a 500 ms exposure with a 250 mW blue LED light source with a 470 nm nominal wavelength (Thorlabs, Newton, NJ, USA) set at 50% intensity, filtered through a standard GFP filter cube set (Chroma, Rockingham, VT, USA). Focus was set for all the frames in the mosaic after daily cell feeding using the mosaic plug-in for Axiovision image acquisition software (Zeiss)

An 18 $\times$ 22 mosaic of 396 contiguous FOVs with a 10% overlap is acquired, covering an area of approximately 180 mm^2. The acquired (x, y, t) data sets over three replicas consist of 529 336 files, and about 1.427 TB of data. Each FOV consisted of (1040×1392) pixels and each time point contained an area of (18×22) FOVs, all FOVs stitched into a
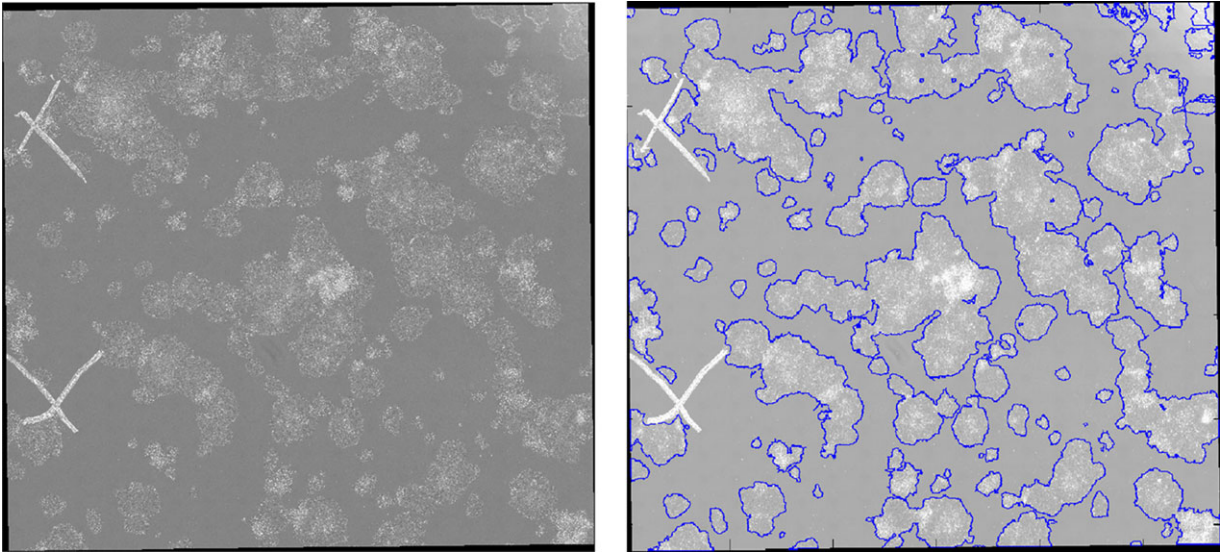
Fig. 4. Segmentation example: Phase contrast image example (left). Corresponding segmented image overlaid on top of the phase image (right).

single time point mosaic consisting of approximately (20 000 × 20 000) pixels. Flat-field images for the GFP channel were collected using a fluorescent solution sandwiched between a glass slide and a coverslip, as previously described (Model & Burkhardt, 2001).

Flat-field and dark current correction

In our work, we are interested primarily in fluorescent microscope image correction models that compensate for thermal noise (dark current models), uneven illumination (vignetting functions) and stem cell media interaction (background correction methods). The dark current and vignetting functions are modelled using recorded single FOV images. We follow a standard calibration procedure described in Eq. (1) that utilizes recorded dark current images acquired with a closed camera shutter and fluorescein images obtained by imaging a reference fluorescent solution (Model & Burkhardt, 2001). In this study, we are not concerned with photo-bleaching, autofluorescence, CCD readout and quantization noise since they represent standalone problems (Wu *et al.*, 2008). The calibration is represented by Eq. (1).

$$GFP_{signal} = \frac{I - B - D}{F - D} = \frac{I - D}{F - D} - B_f, \qquad (1)$$

where I is the intensity of an acquired image, D is the dark image, F is the fluorescein image and B is the background image, the offset information present in I. Note that, B_f is the background image after flat-field correction. Through the rest of the paper all estimation models and analysis are done to measure B_f.

Optimization framework for background correction

Evaluation of background correction error.   For the purpose of evaluating estimated background, we established a set of background reference pixels (BRP) through the following process, which is displayed in Figure 5: The BRP of any given image are the set of pixel locations that belong to the background (BKG) of that image and the foreground (FRG) of the latest image in the given time-sequence dataset ($BRP = FRG_{latest} \cap^B KG$). These are the pixels of known background values but considered as unknown when performing the background correction. When fitting the model we use pixels that are contained in the background BKG but not in the BRP. That is, the set $BKG \cap^B RP^c$ is used for training and the set BRP is used for testing.

Figure 5 shows segmentation of an early time point ($t = 50 \times 45$ min = 37.5 h from the first acquisition on the left). The middle mask shows the background pixels that were considered for model fitting. The third mask shows the background pixels from the latest time point ($t = 161 \times 45$ min = 120.75 h from the first acquisition) that were omitted from the fit and used to compute the remaining error in the background correction.

When applying background correction techniques to the flat-field and dark current corrected GFP images, different results of corrected images are obtained. In order to evaluate multiple background correction techniques, we define an optimization objective function in Eq. (2) with the search space samples. We search for an optimal data partitioning scheme and background correction model that achieve (1) the minimum RMS of the difference between the modelled background intensities and the observed intensities over the BRP (Fig. 5), (2) maximum SNR and (3) minimum execution time (Texec).
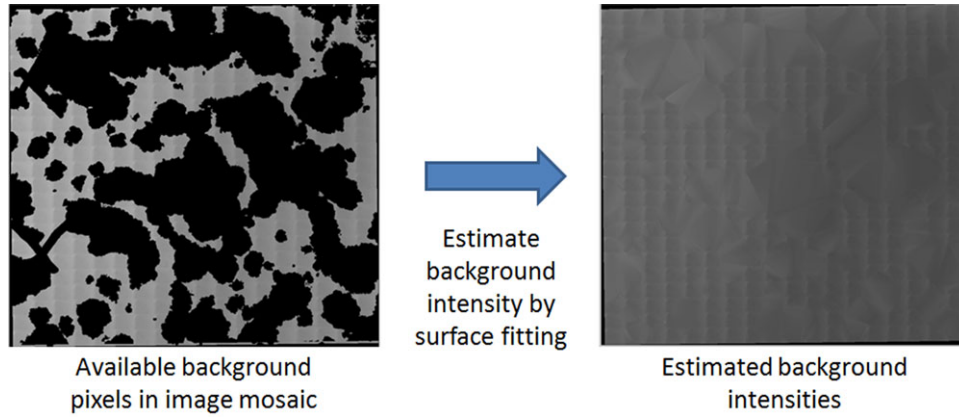
Fig. 5. Three regions of interest: Three regions of interest: Segmented image with all available background pixels labelled as white (left). Segmented image with background pixels used for the fit modelling in white (middle). Segmented image where background reference pixels (BRP) in white were used for evaluating the remaining error as RMS value.

After successful background correction the image SNR is increased by downsampling (replacing a group of neighbouring pixels by their average value). The method that has the maximum SNR after downsampling with the same binning factor of $n \times n$ will be better suited for this type of correction. Each term of the optimization objective function in Eq. (2) is given a weight determined by its importance. In our experiments, the weights followed the ratio $w_{RMS} : w_{SNR} : w_{Texec} = 3 : 2 : 1$. The values of *RMS*, *SNR* and *Texec* were normalized with respect to maximum *RMS* value, maximum *SNR* value, and maximum execution time per time point. All of the terms in the following equations will be described in the following subsections.

Optimization function:

$$\{DataPartition, Model\}^* =$$
$$\arg\min_{\{DataPartition, Model\}} (w_{RMS} * RMS$$
$$+ w_{SNR} * (1 - SNR) + w_{Texec} * Texec). \quad (2)$$

Search space:

$$DataPartition = \{full\,mosaic,\ submosaic\}.$$

$$Model = \{polynomial\,surface,\ cubic\,interpolation,\ linear$$
$$interpolation,\ nearest\,neighbor\}.$$

SNR and the *RMS* error of a given image are computed according to the equations below:

$$SNR = \frac{1}{N_{FRG}} \sum_{n \in FRG} (I_n - \hat{B}_n)/\text{std}\left(\{I_n - \hat{B}_n\}_{n \in BKG}\right), \quad (3)$$

$$RMS = \sqrt{\frac{1}{N_{BRP}} \sum_{n \in BRP} (I_n - \hat{B}_n)^2}, \quad (4)$$

where in Eqs. (3) and (4), N_{FRG} and N_{BRP} denote the number of foreground pixels and BRP, respectively, n is an index for pixel location in the image, I_n is the observed intensity of pixel n, $\hat{B}_n$ is the modelled background intensity of pixel n, FRG is the set of pixel locations that are identified as foreground and BKG is the complement of FRG and contains the set of pixels identified as background, BRP is a subset of BKG used for background model assessment.

Data partitioning schemes and surface fit functions. We have considered two image partitioning schemes, the full mosaic and submosaic-based partitions. Figures 6 and 7 illustrate both data partitioning schemes. Submosaic images in Figure 7 are created by extracting the average value of a sub tile (an example of a subtile is the red square in the upper left corner of a FOV) from each FOV at a fixed location and placing it into a constructed submosaic image according to the FOV index in the grid of FOVs per time frame. Each FOV has a dimension of (1040×1392). For a subtile size of 16×16 pixels, each FOV will be tiled into $65 \times 87 = 5655$ subtiles. Each subtile at a particular location will be replaced by the average value over 16×16 pixels. These values from the same location across all FOVs (the red for example) will be assembled together to form one submosaic. Each submosaic will have a size of (18×22) pixels. The total number of submosaics formed is 5655. The surface fit will be applied on each mosaic independently. Figure 8 shows the submosaic images formed from the (18×22) grid of FOVs. The black pixels indicate that they belong to the foreground and hence are not available for background modelling. The idea behind the submosaic creation is that background intensities are continuous throughout the entire plate and across FOV boundaries. The submosaic image is a good approximated map of the background throughout the entire mosaic representing a particular FOV location.

For the optimization study, we used four surface fitting functions for correcting terabyte-sized GFP images
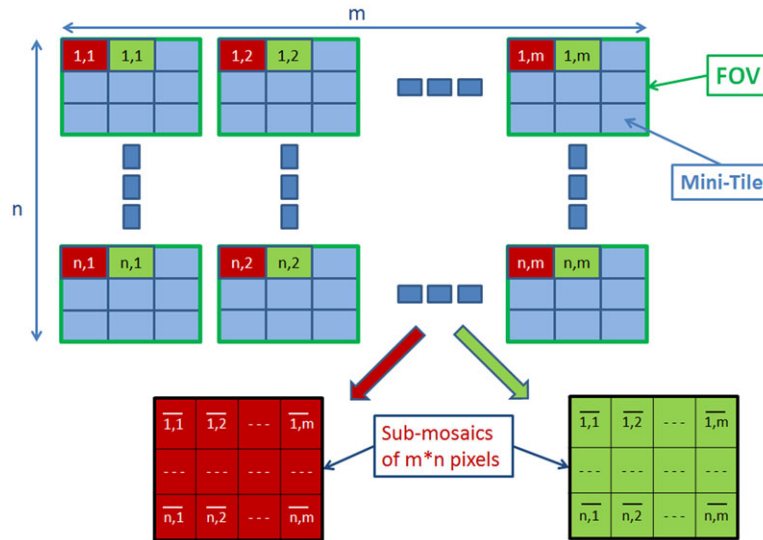
Fig. 6. Image partitioning scheme 1: Background intensity estimation by surface model fitting to an image mosaic based on available background pixels (nonblack pixels in the image on the left).
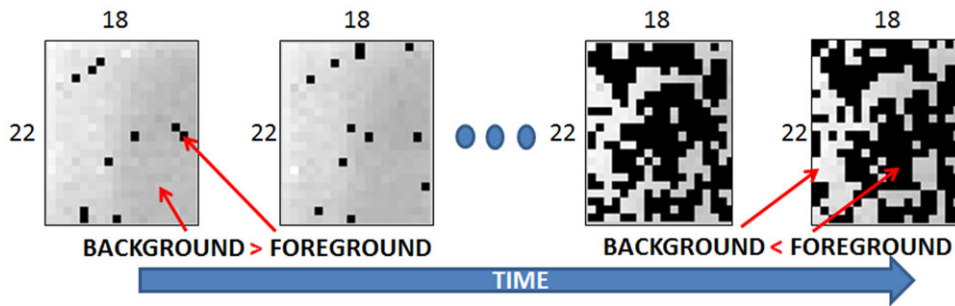

Fig. 7. Image partitioning scheme 2: Average values of the red pixels from the same location of all FOV are assembled into the red submosaic. The same procedure is performed on the pixels at the following location (the green one) etc. until all locations on the FOV have been assembled into submosaics. Background estimation by surface fitting is performed to the constructed submosaics. There are $m = 18$ times $n = 22$ equal to 396 FOVs used for background estimation. The red and green colours denote the original location of the pixels in each FOV and their new locations in a submosaic images.

Table 1. Surface fit functions.

Fit Type	Description
Polynomial surface	$Z = p_{00} + p_{10}*x + p_{01}*y + \ldots + p_{12}*x*y^2 + p_{03}*y^3$
Cubic interpolating spline	This method fits a different cubic polynomial between sets of three points for surfaces.
Linear interpolation	This method fits a different linear polynomial between sets of three points for surfaces.
Nearest neighbour	This method sets the value of an interpolated point to the value of the nearest data point.

as described in Table 1. The methods come directly from the Matlab library (with online documentation found in: http://www.mathworks.com/help/curvefit/fit.html) and are described in detail in Rovenskii (2010).

Results

First, we discuss the results of the best technique for GFP background correction found from our optimization scheme. Then, we compute the best image mosaic downsampling to get the SNR to a minimum value of 5 per Rose criterion (Watts *et al.*,

2000) across all time points and replicas, which gives a good SNR whereas minimal spatial resolution is sacrificed

Optimization results

We report optimization results performed on 15 time points in Figure 9 below. The time-points are selected by increment of 10 frames. The execution time of the model fit is measured in seconds and it is the time needed to perform the background
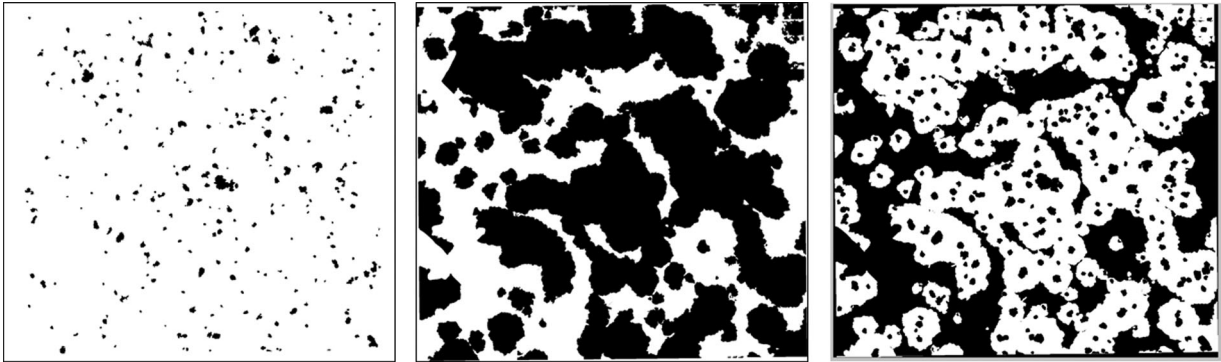
Fig. 8. Resulting submosaic images over time: The decreasing number of background pixels is due to the growth of stem cell colonies. The submosaic images on the left are at the early time points and on the right at the later time points of the experiment. Nonblack pixels in 18×22 submosaic images are at those FOV locations that contain only background pixels within a FOV. The black pixels in 18×22 submosaic images represent the average intensity value of the foreground pixels in that particular FOV.
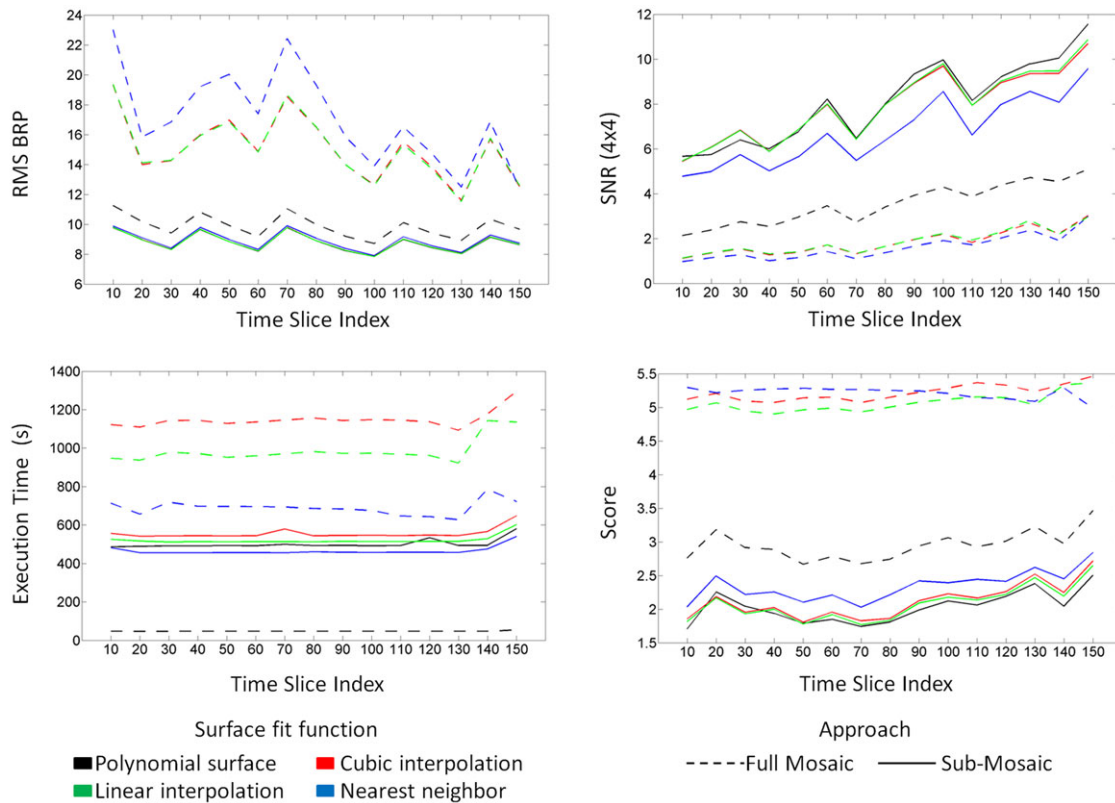

Fig. 9. Results between image partitioning schemes: This figure displays the results for image partitioning approaches for 15 time-slices (one slice is taken every 10 frames). The four illustrations above represent the quality metrics for the eight partitioning-model combinations (two image partitioning schemes and four background interpolation models performed on 4×4 downsampled GFP images to meet the SNR constraint). The 'score' quality metric is computed according to Eq. (2) with the *RMS*, *BRP*, *SNR* and *Texec* normalized by the maximum values over both approaches.

estimation on a mosaic image at one time point. The overall execution time including reading image tiles, model fit, background correction and mosaic assembly takes about 15 min per time point. The optimization code was written in Matlab® [version 8.0.0.783 (R2012b)] with the additional packages 'image processing toolbox 8.1' and 'statistics toolbox 8.1'. The

code was executed on a desktop machine with Intel® Xeon® CPU E5–2620 0 @ 2.00GHz, 64.0 GB RAM, and Windows 7 Enterprise (64 bit) operating system.

The results displayed in Figure 9 indicate that for each of the four considered surface fit functions (see Table 1), the approach based on fitting submosaic images achieves smaller

scores, smaller RMS BRP and at least twice larger SNR than the approach based on full mosaic images. Figure 9 also shows that the polynomial surface background model is better than the rest of the fitted models in either data partitioning scheme. One can observe that the differences between the fitted models are much smaller within the submosaic data partitioning approach than within the full mosaic approach. Even though the images do not appear the same visually, the optimization entries (RMS level of noise, SNR after downsampling and execution time) are very similar. Overall, using the submosaic-based data partition and a polynomial surface background model produced the best GFP background correction, according to our optimization score. We will use this technique to correct the GFP background for all time points and replicas.

To highlight the effectiveness of this new approach, we arbitrarily selected frame 50 and corrected it by assuming that the background is constant through the entire mosaic. Hence, one background image (location 4 in Fig. 1) is taken as the background for all tiles in the mosaic. The resulting image gave a SNR value of 3.150 and an RMS value of 15.33. This illustrates that a traditional approach is not a good solution to solve the background correction problem in a large mosaic dataset.

Spatial downsampling and signal-to-noise ratio constraint

Due to the low SNR present in the dataset for all time points, we minimize the spatial downsampling ratio (maximizing spatial image resolution) while satisfying the Rose detectability criterion (Watts *et al.*, 2000) for the SNR over all image time frames and downsampling ratios according to Eq. (5) and for all replicas. The Rose criterion of SNR > 5 is met for the kernel sizes of (4 × 4) and higher across all temporal frames and replicas (Fig. 10A). Figure 10(B) shows the normalized histograms of the background and foreground signals for the same image and time point as the one displayed in Figure 2. This highlights the contrast between the (1 × 1) downsampling with low SNR and the (4 × 4) downsampling with higher SNR.

$$DownsamplingRatio^* = \min_{SNR(DownsamplingRatio, data(x,time))>5}$$
$$Downsampling\,Ratio, \quad (5)$$

$$DownsamplingRatio = \{2 \times 2, 4 \times 4, 6 \times 6, 8 \times 8, 10 \times 10,$$
$$12 \times 12, 14 \times 14, 16 \times 16\}.$$

Figure 11 shows examples of the results of three time point mosaics corrected using submosaic and the polynomial fit and binned at (4 × 4).

Sensitivity analysis

In order to analyse the background correction performance for increasing colony coverage, we compared results at time

points with small and large colony coverages. First, we selected an image taken at an early time point (frame number 20) where colonies are still very small and their coverage is only 3.22% of the image. SNR values are always computed at the time point 20. Next, we took segmented images from time points 40 until the end of the experiment with an increment of 20 frames. The purpose is to compute the measures such as the RMS of the BRP pixels, the SNR at 1 × 1 computed on top of the colonies at time point 20, and the SNR at 4 × 4 after the image is binned. We computed these values after each correction as a function of the percent colony (foreground) coverage. The colonies in the last time point (frame 160) covered only 58% of the plate. We had to dilate the segmented mask of that last point in order to cover 70%, 80% and even 90% of the plate.

Figure 12 shows the results of the background correction performance for increasing colony coverage. As colonies grow and cover more background, the accuracy of estimating the colony signal computed at time point 20 drops significantly even after doubling the colony area from 3.2% to 7%. This increase in colony coverage translates to a significant drop of the SNR. The SNR drop reflects an increased uncertainty of the estimated colony signal at later time points. The uncertainty increases radially in each colony towards the colony centre as the values are farther from the known background pixels. Nevertheless, the RMS value is almost unchanged with an increased colony coverage which indicates a constant noise level in the image.

Discussion

We discuss an additional evaluation aspect of background correction with respect to measured dark current noise. We view the measured dark current noise as the lowest possible noise level in our measurement system and compare the remaining noise after background correction to the dark current noise.

The dark image subtracted from the fluorescein and the current image I in Eq. (1) is a median image computed over all 396 dark images that are acquired with closed shutters. However when an image is being acquired it has its own corresponding dark, background and signal information. Thus, Eq. (1) can be written as

$$\frac{I - D}{F - D} = \frac{(S + B + D_i) - D}{F - D}. \quad (6)$$

Assuming no signal or background: $S = 0$ and $B = 0$, then Eq. (6) becomes

$$\frac{I - D}{F - D} = \frac{D_i - D}{F - D}. \quad (7)$$

The RMS value computed on the image by Eq. (7) is the lowest white noise level in our system. Figure 13 shows the RMS value computed over 100 dark images following Eq. (7). The fluctuation of the dark noise is approximately 5.07 whereas the total RMS value for the GFP correction including the dark, fluorescein, and background correction is on
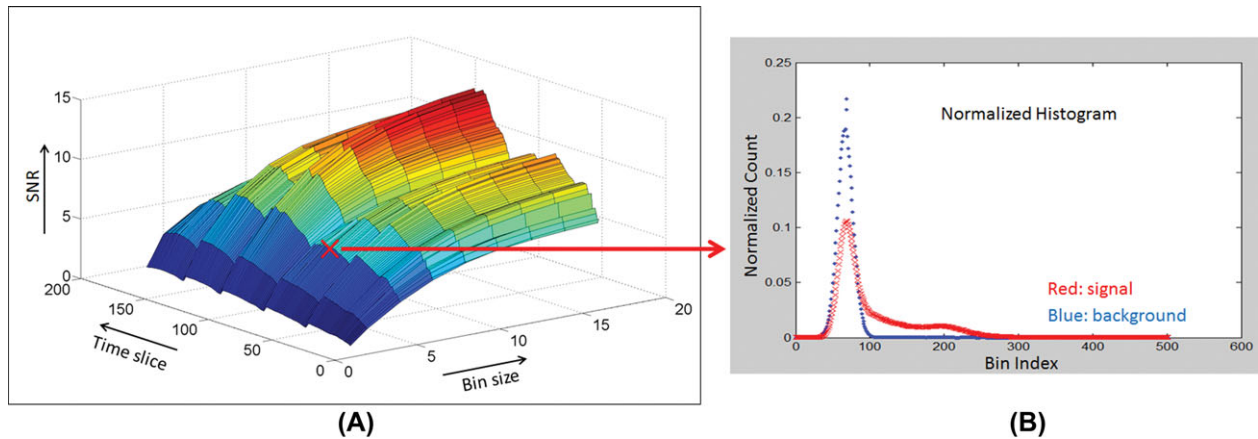
Fig. 10. Downsampling and SNR plot: (A) SNR computed over all 161 temporal frames of replica 1 and spatial downsampling kernels between 2×2 and 16×16. The Rose criterion of SNR > 5 is met for the kernel sizes of 4×4 and higher across all frames. (B) Foreground and background normalized histograms of one binned image (4×4) with 500 bins. The two curves illustrate a SNR $= 6.1$ in contrast with the SNR $= 1.4$ in the (1×1) binned image.
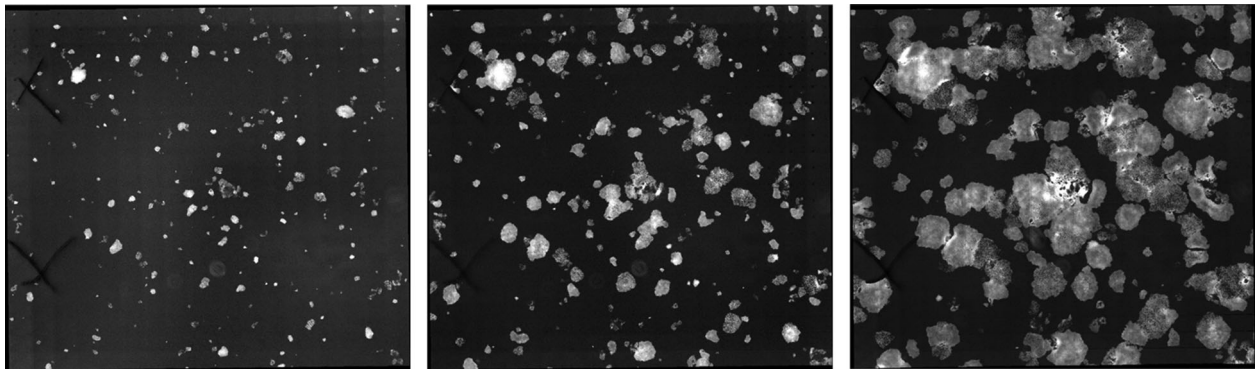

Fig. 11. Examples of corrected GFP images: Corrected images for frames 50, 100 and 150 from left to right. These images are spatially downsampled by (4×4), flat-field correction, dark current correction and background subtraction.
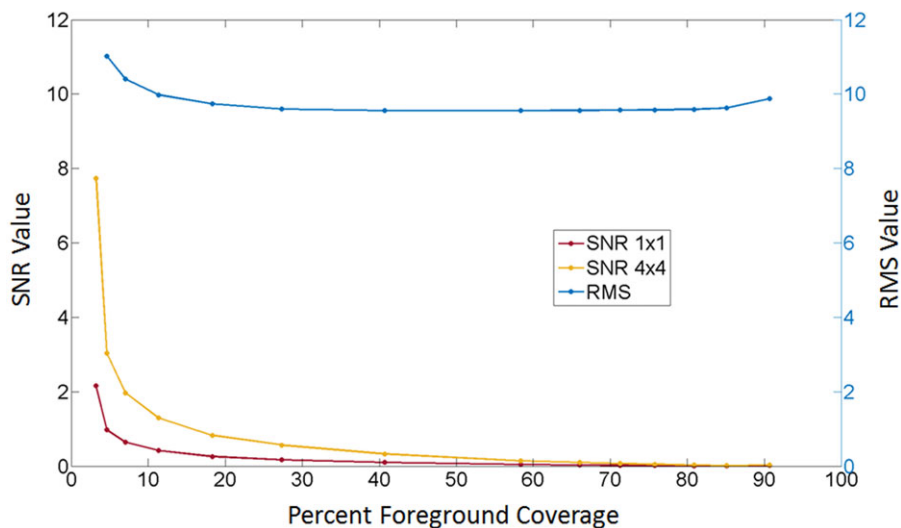

Fig. 12. Sensitivity analysis to the foreground percentile coverage: Image frame 20 is considered for this sensitivity analysis where colonies are still very small and cover only 3.22% of the image. SNR values are computed at the image frame 20. RMS values are computed over known background pixels that are considered unknown by taking segmented images from image frame 40 until the end of the experiment with an increment of 20 frames. The colonies at the end of the experiment cover 58%, morphological dilation is done to get the colony coverage to 90%.
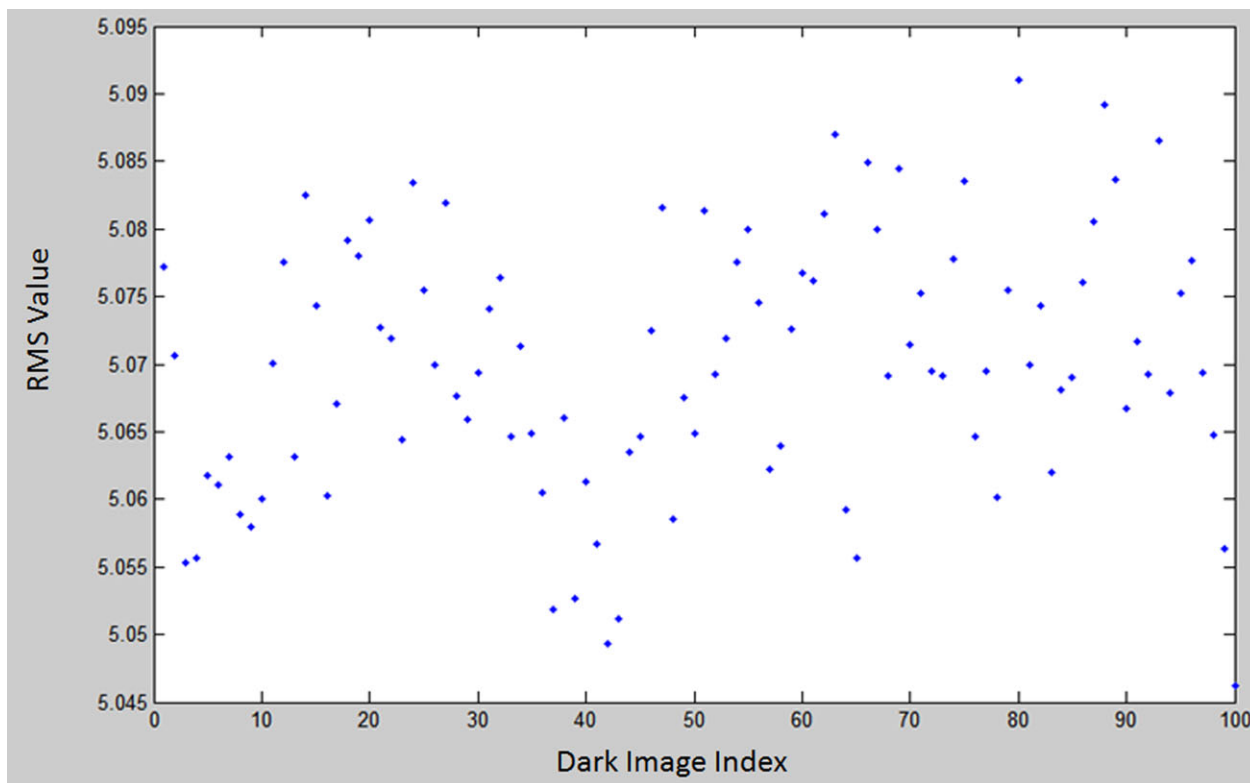
Fig. 13. Dark images error: RMS error obtained over 100 single FOV images with only background pixels.

the order of 8.9. That gives a maximum error contribution of $e = 9.53 - 5.07 = 3.83$ as total error from fluorescein and background correction. This error is less than the dark noise error contribution which demonstrates experimentally that the background noise contributing to the total RMS value can be corrected computationally with the developed method to a magnitude less than the measured dark current noise.

Conclusions

We demonstrated a computational technique for background correction of terabyte-sized fluorescent microscopy images. The solution was obtained by optimizing over a space of two image partitioning schemes and four background correction models. The optimization cost function included RMS, SNR and execution time. We concluded that the optimal solution framework contribution from background noise to the total RMS value does not exceed the measured dark current noise. The new technique generated a corrected image with half of the error value and twice of the SNR value when compared to an approach that assumes constant background throughout the mosaic. We made the code is available as open-source from the following link https://isg.nist.gov/.

In the future, we plan to develop a temporal model of colony background which would include photo-bleaching effects, movement of the media, illumination variation, and other variables in order to minimize the uncertainty of estimated colony pixels, especially for those time points with large colony coverage.

Disclaimer

Commercial products are identified in this document in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the products identified are necessarily the best available for the purpose.

References

Chalfoun, J., Kociolek, M., Dima, A., Halter, M., Cardone, A., Peskin, A., Bajcsy, P. & Brady, M. (2013) Segmenting time-lapse phase contrast images of adjacent NIH 3T3 cells. *J. Microsc.* **249**(1), 41–52.

doi:10.1111/j.1365-2818.2012.03678.x. http://www.ncbi.nlm.nih.gov/pubmed/23126432.

Chalfoun, J., Majurski, M., Peskin, A., Breen, C. & Bajcsy, P. (2014) Empirical gradient threshold technique for automated segmentation across image modalities and cell lines. *J. Microsc.* (under review), 1–18.

Chen, G., Gulbranson, D.R., Hou, Z. *et al.* (2011) Chemically defined conditions for human iPSC derivation and culture. *Nat. Meth.* **8**(5), 424–429. doi:10.1038/nmeth.1593. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3084903&tool=pmcentrez&rendertype=abstract.

Dice, L.R. (1945) Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302.

Galego, R.M.F. (2011) Geometric and radiometric calibration for pan-tilt surveillance cameras. *Universidade T´ecnica de Lisboa Instituto Superior Te´cnico* **57**. https://dspace.ist.utl.pt/bitstream/2295/990125/1/dissertacao.pdf.

Goldman, D.B. (2010) Vignette and exposure calibration and compensation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(12), 2276–2288. doi:10.1109/TPAMI.2010.55. http://www.ncbi.nlm.nih.gov/pubmed/20975123.

Kim, S.J. & Pollefeys, M. (2008) Robust radiometric calibration and vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(4), 562–576. doi:10.1109/TPAMI.2007.70732.

Leong, F.J.W-.M. (2003) Correction of uneven illumination (vignetting) in digital microscopy images. *J. Clin. Pathol.* **56**(8), 619–621. doi:10.1136/jcp.56.8.619. http://jcp.bmj.com/cgi/doi/10.1136/jcp.56.8.619.

Model, M.A. & Burkhardt, J.K. (2001) A standard for calibration and shading correction of a fluorescence microscope. *Cytometry* **44**(4), 309–316. http://www.ncbi.nlm.nih.gov/pubmed/11500847.

Piccinini, F., Bevilacqua, A., Smith, K. & Horvath, P. (2013) Vignetting and photo-bleaching correction in automated fluorescence microscopy from an array of overlapping images. *Researchgate.net*, 464–467. doi:10.1109/ISBI.2013.6556512. http://www.researchgate.net/publication/233779359_Vignetting_and_photo-bleaching_correction_in_automated_fluorescence_microscopy_from_an_array_of_overlapping_images/file/9fcfd5134f61f6dcbc.pdf.

Rovenskii, V.Y. (2010) *Modeling of Curves and Surfaces with MATLAB. Labkom.stikom.edu.* Springer, New York. http://labkom.stikom.edu/download/ebook/Mathematic Modern/1461401216.pdf.

Saha, K. & Jaenisch, R. (2009) Technical challenges in using human induced pluripotent stem cells to model disease. *Cell Stem Cell* **4**(5(6)), 584–595. doi:10.1016/j.stem.2009.11.009.

VanDenBerg, D.L.C., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J. & Poot, R.A. (2010) An Oct-4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**(4), 369–381. http://www.ncbi.nlm.nih.gov/pubmed/20362541.

Waters, J.C. & Swedlow, J.R. (2007) Techniques interpreting fluorescence microscopy images and measurements. In *Evaluating Techniques in Biochemical Research* (ed. by D. Zuk), pp. 36–42. Cell Press, Cambridge, MA. http://www.cellpress.com/misc/page?page=ETBR.

Watts, R., Wang, Y., Winchester, P.A., Khilnani, N. & Yu, L. (2000) Rose model in mri: noise limitation on spatial resolution and implications for contrast enhanced MR angiography. *Intl. Soc. Mag. Reson. Med.* **4**(8), 462.

Wu, Q., Merchant, F. & Castleman, K. (2008) *Microscope Image Processing.* Academic Press; Elsevier, Burlington, MA.

Zwaka, T.P. & Thomson, J.A. (2003) Homologous recombination in human embryonic stem cells. *Nat. Biotechnol.* **21**(3), 319–321. doi:10.1038/nbt788.