

# Minimizing Image Artifacts in Microarray Imaging

By Kevin Ryan - Media Cybernetics  
March, 2002

## Abstract

The process of imaging a microarray sample can introduce a great deal of noise and bias into the data. The results can have a higher variance than the original signal, swamping the useful information.

This paper presents a general framework for determining the sources of imaging distortion, and a set of techniques for improving data quality. Poisson noise is discussed, as well as offset and scaling biases. In addition, the error introduced by image segmentation techniques is considered, leading to the recommendation that segmented spot outlines not be used in determining intensity measurements. Image processing useful for determining microarray geometry and positioning is compared to image processing appropriate for measuring spot intensities

Finally, there is a brief discussion of tracking the imaging process for consistency and identifying outlier or invalid data.

## Artifact Identification

Table 1 lists error and artifact sources for microarray acquisition systems based on CCD cameras.

Note that scanner based systems will have essentially the same problems, although linear scanner pixel defects will affect entire columns of the image, and illumination intensity is likely to be more uniform in a scanner.

Pixel defects are an inherent problem for any solid state detector – there may be pixels that will simply not respond to illumination properly. If a manufacturer's pixel defect map is available, suspect pixels or columns will be generally replaced by the values of adjacent pixels. The general effect is to reduce the spatial sampling of the image, with a zero-bias effect on integrated object intensity. If pixel defect correction is not done, the pixels will produce a non-zero bias to the affected objects.

Dark current is the thermal noise accumulated in each pixel. This can be estimated from multiple no-light exposures of different lengths, and subtracted from a sample. It is best avoided with short exposures or cooled detectors.

Readout noise is Poisson statistic random noise in counting the intensity accumulated in a pixel, due to the electronics involved. It does not scale with exposure time, and therefore is best reduced by high signal values.

Pixels will also tend to have a fixed bias, which will be present even at zero exposure times. In practice it is measured along with the readout noise, as an estimated per-pixel constant bias to any image. This bias image will be the remainder after removing exposure scaled dark currents.

Flat fielding is particularly troublesome in microscope imaging, although it is also an issue in scanners. It can be measured by imaging uniform bright samples to determine what the intensity response is across the imaging field of view. The uniform sample might be a fluorescent glass, a uniformly stained slide, or any other featureless object. In microscopy this is often done with a defocused object, in order to reduce any remaining sample

**Table 1 - Imaging artifacts**

<b>Artifact</b>	<b>Source and Characteristics</b>	<b>Measurement</b>	<b>Correction</b>
Pixel Defects	Individual pixel errors, sometimes columns or clumps	Manufacturers Defect Map; if not available examine image statistics for zero or non-linear pixel responses	Replace with nearby pixels during capture - performed by many camera drivers
Readout/ Amplifier Noise	Poisson statistic amplifier error, periodic EMF	Estimate from multiple images of the same exposure time. Periodic noise can be found with Fourier analysis	Periodic EMF implies a need for better shielding and/or isolation. Amplifier errors can be averaged to a constant background with known variance; subtract from images
Dark Current	Thermal noise, dependent on integration time	Series of integrations in no-light conditions to establish relationship of exposure to thermal noise, minus readout noise	Subtract dark current scaled by exposure time from image data
Pixel Bias	Base unexposed pixel value, inherent in the detector, relatively stable.	Calculate dark current - readout average; bias is the remaining pixel intensities	Subtract bias image from sample. This can be combined with the Readout noise correction
Non-flat Field	Position dependent strength of illumination, camera response, optical defects	Image bright evenly stained samples, or interpolate from array of uniform dots, stained microbeads, or other uniform samples, minus dark field	Divide offset corrected data by flat field; this can be scaled back into range with a multiplicative constant
Imaging Background	Mounting material, ambient light and scatter	Average images of unstained mounts, minus dark current and readout noise; correct for shading. This may be further smoothed to improve the estimates	Subtract from shading corrected image
Local Sample Background	Non-specific staining, incomplete washes.	Estimate background at a distance from located spots; if spot position varies greatly, use nearby backgrounds for 'negative' spots to avoid classifying low value negatives as background	Subtract from local spots; flag spots with high local backgrounds as dubious
Segmentation Errors	Linear and non-linear errors in locating the edges of sample spots. Non-zero bias: tends to underestimate integrated object intensities by rejecting object edges.	Image sets of spots labeled at different intensities, examine variances for each labeling intensity	Limit segmentation to locating grids; calculate spot intensity over larger area minus average background

structure. This can compensate for both illumination and relative pixel sensitivities. Note that failing to perform flat fielding may result in order of magnitude differences in object intensities due to positioning.

Imaging background is a product of the acquisition system and the sample presentation. It can be affected by mounting media, sample washes, and other effects. It can be estimated by repeatedly imaging blank treated media, and averaging the results to obtain a background estimate. Note that these images must be corrected prior to averaging for camera noise, camera bias, and illumination shading.

Local sample backgrounds are more difficult to measure; they are a product of the sample preparation. They can be estimated by examining the region immediately around a deposited spot, but this estimate will greatly add to uncertainties in the measured object intensities. This is particularly true if the local background is due to non-specific binding problems, as the binding in the spot may displace some of the background, adding further bias to the local background estimate. Spots with high local backgrounds should be considered suspect despite correction.

Segmentation errors are an inherent product of using image processing to locate objects of interest. Note that this is an error with a non-zero bias. Avoiding segmentation problems is discussed in some detail later in this poster.

### Correction Sequence

Determine the thermal noise on a pixel by pixel basis,  $Noise_t$ . Also determine the combined pixel bias and readout noise,  $Bias_r$ . Given a sample image  $Sample$ , with exposure  $T_{exposure}$ , the detector corrected image  $S_d$  is:

$$Image_{detector} = Sample - Bias_r - (Noise_t * T_{exposure})$$

Next, correct for illumination errors  $Shading$  to obtain a uniform response sample corrected for the imaging system:

$$Image_{uniform} = Image_{detector} / Shading * avg(Shading)$$

The multiplication by the average of the shading is optional, but useful to keep the signal in the same basic dynamic range. At this time sample backgrounds can be removed, using the corrected global background  $Background_g$ .

$$Image_{corrected} = Image_{uniform} - Bias_{background}$$

Further correction can be obtained by estimating the local backgrounds, however this correction is best applied to individual spots rather than the image as a whole.

### Signal to Noise

The following discussion is taken from Sky and Telescope (1).

For a single pixel, the signal is the sum of the object signal, the dark current, pixel bias, global and local backgrounds. The noise is calculated from the noise in all these components, as well as the readout noise:

$$S = S_{obj} + S_{dark} + S_{bias} + S_{bkgr}$$

$$N^2 = N_{obj}^2 + N_{dark}^2 + N_{bkgr}^2 + N_{rdout}^2$$

$$S/N = \frac{S_{obj} + S_{dark} + S_{bias} + S_{bkgr}}{\sqrt{(N_{obj}^2 + N_{dark}^2 + N_{bkgr}^2 + N_{rdout}^2)}}$$

Taking advantage of the square root nature of the noise (except for readout noise, which cannot be simplified since it's not a photon counting phenomenon), we can replace some terms. For a single pixel in the camera:

$$S/N = \frac{S_{obj} + S_{dark} + S_{bias} + S_{bkgr}}{\sqrt{(S_{obj} + S_{dark} + S_{bias} + N_{rdout}^2)}}$$

Assuming bias, dark current, and flat field images with high S/N and low uncertainty, as well as a good background estimate, we can derive a simple expression for object S/N.

We define  $n$  as the number of pixels,  $C_{object}$  as the integrated object intensity,  $C_{back}$  as the integrated global background and camera bias, and  $N_{readout}$  as the readout noise per pixel:

$$S/N = \frac{\sqrt{C_{object}}}{\sqrt{\left\{1 + \frac{C_{back}}{C_{object}} + n * \frac{N_{readout}^2}{C_{object}}\right\}}}$$

To improve the total S/N, there are several approaches

- Minimize the background count, using cooled detectors and clean sample preparations.
- Use detectors with low readout noise.
- Maximize pixel size, as readout noise is proportional to the number of pixels, not their size.
- Minimize the pixel count, reducing object magnification to cover as few pixels as possible.
- Minimize exposure time to reduce thermal noise.
- Maximize the integrated object value.
- Minimize local backgrounds in sample preparation.

Note that local backgrounds are particularly susceptible to error, as they are highly non-uniform sample problems.

### Inherent Segmentation Errors

An additional source of error in object intensity estimation can come from image segmentation. Image segmentation for object location is generally based on intensity thresholds, with all content within some thresholded perimeter classified as an object. Pixels within this perimeter that are not stained will be incorrectly classified as object, while edge pixels that do not pass the threshold will be incorrectly classified as background.

In the case of microarray scanning, where the total staining of the object is the primary value of interest, segmentation errors will tend to lower the integrated intensity. This means that segmentation errors will add error with a negative bias. Due to the difficulties in thresholding dim objects, this error will increase for low intensity

spots. In the extreme case, where the spot outlines cannot be identified, placing an object outline in an estimated position, the error will be extremely large.

In addition, the errors induced by segmentation will be dependent on the segmentation algorithm used. If neighborhood filtering, local contrast, or edge detection methods are used, the segmentation estimate becomes non-linear with respect to spot intensities.

**For these reasons, it is recommended that microarray spot intensities be estimated by integrating over a larger area encompassing the spot, rather than from segmented outline areas.**

Object segmentation can provide for the orientation of sampling grids for microarray spots, and for tracking sample problems. Integration area should be a compromise between an area large enough to cover segmentation and spot placement variations on the one hand, and minimizing the integrated background on the other.

## Quality Tracking

Imaging of microarrays provides a number of features that can be tracked for quality assurance:

- Spot shapes
- Standard deviations of spot intensity, both within and between spots
- Intensity of standards samples
- Global backgrounds for sample quality
- Local backgrounds for spot quality
- Background standard deviations
- Position variation for spots: this reflects on deposition and substrate variations
- Flat field changes, measured from uniform samples

Note that segmentation features, such as spot shape and position variations, should be weighted by spot intensity, as low intensity spots will have higher segmentation errors.

Staining chemistry will also play a role here, as some methods result in extremely uneven spotting. Setting acceptable thresholds for quality features will be a matter for the experimenter, but it is suggested that the calculated S/N for integrated spot intensity be a part of this determination.

Tracking backgrounds and flat fields is particularly important in microscopic systems, as microscopes are susceptible to lamp aging, environmental lighting, and the urges of experimenters to adjust the illumination. Failure to track illumination can result in the invalidation of a considerable number of experiments.

## Conclusions

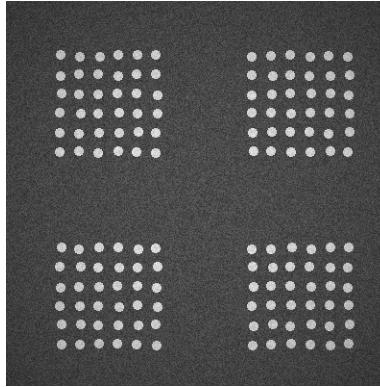
Quantitative image analysis requires tracking and retaining a considerable amount of information, due to the number of parallel detectors in use – but the correction procedures are relatively straightforward.

The best results can be obtained by thoroughly measuring the imaging system, providing the best samples possible, and avoiding overly precise segmentation. Combined with tracking the performance of your imaging system, simple techniques can greatly improve the accuracy of your results.

## Appendix A: Images and Results

Below are some images demonstrating the improvement in data from appropriate image correction. These were generated with known distortions, noise, shading, etc., for the purpose of demonstration.

First, an image acquired with random noise, background offset, and illumination shading:



**Figure 1 - Raw Image**

A shading reference image (Fig. 2) and a background image (not shown) were acquired using the same parameters. The shading image is that of a uniform object, revealing the illumination variations across the field:



**Figure 2 - Shading**

Fig. 3 shows the corrected image. Note that the spots are now uniform, as expected:

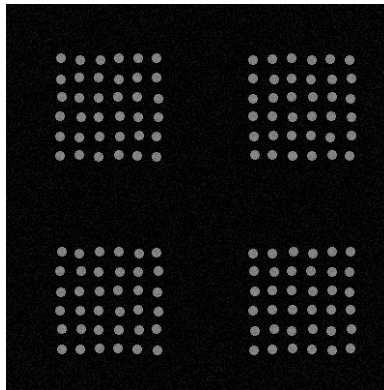


Figure 3 - Corrected Image

Analysis of the images shows the improvements based upon image correction. The raw data is taken from the original image, the background rings are 3 pixel wide rings outside the spots, and the corrected data shows the improvement possible.

Correction consisted of:

$$\text{Corr} = (\text{Raw} - S_{bkgr}) / (\text{Shade} - S_{bkgr}) * \text{Mean}_{shade}$$

Integrating the array spot densities was corrected with the background values, scaled by the spot area:

$$\text{IOD}_{corr} = \text{IOD}_{raw} - (\text{Mean}_{bkRing} * \text{Area})$$

The following table shows the results.

Table 2 – Correction Results

	Mean	Std. Deviation
Raw Data	33529	1452
Background Rings	20587	916
Corrected Data	12942	789

Note that the standard deviation of the spots goes down considerably when properly corrected, leaving only variations due to acquisition noise.

---

**FootNote:**

- (1) S/N derivation from Sky and Telescope; “The Signal to Noise Connection”,  
<http://www.skypub.com/imaging/ccd/signalnoise.html>

**Useful References:**

- (1) Video Microscopy : The Fundamentals (2<sup>nd</sup> Ed.), 1997 – Shinya Inoue, Kenneth R. Spring
- (2) Fluorescence Microscopy of Living Cells in Culture, Part B: Quantitative Fluorescence Microscopy – Imaging and Spectroscopy (Methods in Cell Biology), 1990 - D. Lansing Taylor, Yu-Li Wang (Editor)