

Noise and Colocalization in Fluorescence Microscopy: Solving a Problem

Jeremy Adler,¹ Fredrik Bergholm,² Stamatis N Pagakis,³ and Ingela Parmryd.¹

1. Cell Biology, The Wenner-Gren Institute, Stockholm University, Sweden. 2. Dept. of Numerical Analysis and Computer Sciences, KTH, Stockholm, Sweden. 3. Biological Imaging Unit, Biomedical Research Foundation Academy of Athens, Athens, Greece.

BIOGRAPHY

Jeremy Adler obtained a PhD from Edinburgh University in physiology and pharmacology and subsequently worked in microscopy and image analysis at Nottingham University and the National Institute for Medical Research (London). He is currently working at the Wenner-Gren Institute in Stockholm on image analysis and models of the plasma membrane.



ABSTRACT

The Pearson correlation coefficient (PCC) is regularly used in colocalization measurements, but it is sensitive to image noise. Images of fluorophores are usually degraded by Poisson and background noise and we have found that, even with apparently high quality images, the measured PCC is substantially understated, to the extent that the numbers become misleading. This means that ostensibly significant differences in the PCC between two populations may just reflect differing image quality while dissimilar levels of noise may mask significant differences. A new correction, based on measurements of image quality, derived from a pair of images for each fluorophore, aligns the measured PCC with the true PCC. Our method, the Replicate Based Noise Corrected Correlation (RBNCC), generates an accurate PCC even from poor images. It is highly photon efficient and therefore well suited for use in live cell fluorescence imaging or with rapidly bleaching fluorophores.

KEYWORDS

confocal microscopy, fluorescence microscopy, image noise, colocalization, Pearson correlation coefficient

AUTHOR DETAILS

Dr Jeremy Adler,
Cell Biology,
The Wenner-Gren Institute,
Stockholm University,
106 91 Stockholm, Sweden
Tel: +46 (0)8 16 2759
Email: adler.jeremy@cellbio.su.se
www.wgi.su.se www.colocalize.com

Microscopy and Analysis 22(5):7-10 (EU), 2008

INTRODUCTION

In multiprobe fluorescence microscopy fluorophores are often described as being colocalized. Usually this means that they appear to reside in the same cellular compartment(s), which might be a trivial observation, perhaps due to a physicochemical similarity. A more interesting alternative is that the degree of colocalization varies with the physiological state and there is an underlying molecular interaction. The difficulty lies in quantifying the colocalization. Importantly the accuracy and precision of any quantitation determines the minimum size of a detectable biological response.

If there is a molecular interaction, either direct or indirect, then a local variation in the concentration of one molecule will be matched, to some degree, by a similar or possibly inverse variation in the concentration of the other molecule. This is difficult to detect by eye because our visual skills do not extend to the accurate decoding of pixel intensities, let alone to the making of comparisons between pixels in a pair of images. Pixel by pixel comparisons are often shown using overlay pictures (Figure 1) which are made by inserting separate images into the red and green layers of an RGB image, producing yellow pixels when the intensities of the fluorophores are similar. However overlays can be misleading,

since changes to the detector gain, or fiddling in Photoshop, alter which pixels appear to have matching intensities: so there is a lack of objectivity. An alternative is to use a scattergram (Figure 1), which shows the frequency with which combinations of intensities from homologous pixels occur and reveals any overall relationship between the variation in intensity of the two fluorophores. The disadvantage of this pixel-by-pixel comparison is that the connection between both neighbouring pixels and to cellular structures is lost, though scattergrams from a region of interest (ROI) can also be examined separately.

Colocalization clearly requires quantitation and Manders [1] introduced the Pearson correlation coefficient (PCC) to fluorescence microscopy:

$$r = \frac{\sum (R_i - R_{av})(G_i - G_{av})}{\sqrt{\sum (R_i - R_{av})^2 \cdot \sum (G_i - G_{av})^2}}$$

The numerator is the pixel-by-pixel sum of the product of the intensities of the two fluorophores (R_i and G_i) minus their mean (R_{av} and G_{av}) over the whole dataset. For a single pixel the product can be positive, both intensities are either above or below their mean, or negative, when one is above and the other below their mean. The sum, positive or negative, will be maximised when in individual pix-

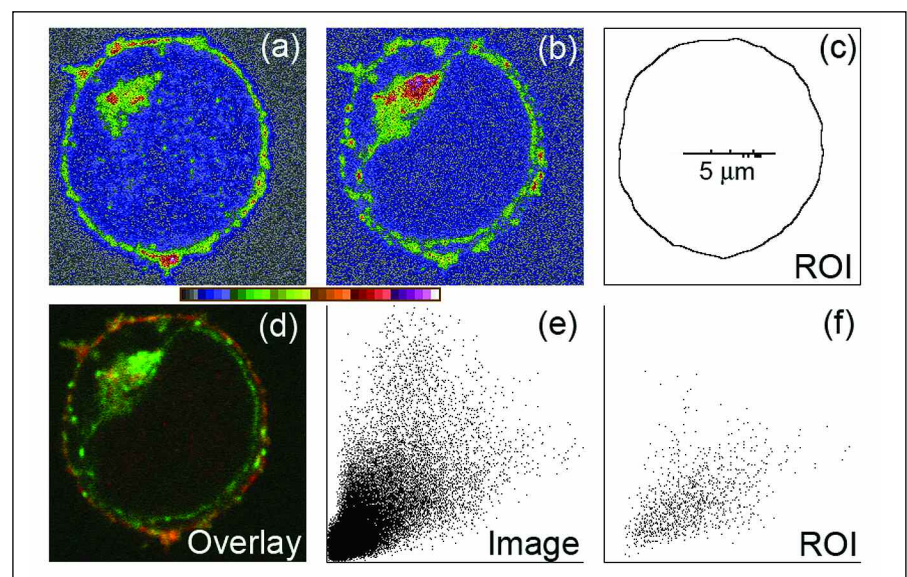


Figure 1:

Standard display of colocalization. (a) and (b) are false colour images of different fluorophores. A region of interest marking the plasma membrane (c). An overlay images is shown (d) with (a) inserted into the green and (b) inserted into the red layer of an RGB image. Scattergrams showing the combinations of intensities found in homologous pixels are shown of (e) the whole image and (f) the ROI. The fluorophores show Lck and CD3.

els both fluorophores have similar relative departures from their respective means. A further feature of the PCC is that intensities distant from the mean carry more weight than those closer to the mean [2].

The denominator appears similar to the numerator but actually calculates the maximum possible product, if the sets of data were perfectly correlated. The division of the numerator by the denominator therefore gives the PCC a range from 1 (a perfect correlation) through 0 (no correlation) to -1 (a perfect but inverse correlation). Importantly the PCC is independent of the gain and offset (two very frequently adjusted parameters during image acquisition), provided the data remains within the detection range.

While the equation is quite simple, its simplicity includes the very basic assumption that the raw data is free from noise (Table 1), which may not hold in fluorescence microscopy.

THE PROBLEM

Ideally an image of fluorescence would record the location and number of fluorophore molecules. However images are based on the detection of a small fraction of the emitted photons that pass through optical filters and are brought to focus on an imperfect detector. The problem starts with the emission of photons: fluorescence is stochastic and the interval between emissions is variable. As a consequence, when two sequential images are acquired of the same specimen there will be differences between the intensities of homologous pixels. This type of noise is called Poisson noise and the variability in the number of photons emitted depends upon the square root of the long term mean: with a true count of 100 photons, an individual reading will probably (67% of the time) fall within the range 90-110. In practice our expensive microscopes do not provide either the photon count or even an estimate of the photon count, but 100 photons per pixel is quite a respectable count for a line-scanning confocal microscope [3]. Our basic problem is that we don't know the number of molecules or even the number of photons in each pixel but instead have a single sample of an intrinsically unstable signal. The fluorescent signal is also affected by the stability of the illumination source and the quantum efficiency of the detector.

Background noise also reduces the quality of images and this category includes events that are unconnected and therefore uncorrelated with the fluorescent signal.

When we started to make measurements of colocalization using the PCC we quickly realized that there was a connection between the measured PCC and the quality of the images. Clearly if the images showing each fluorophore do not accurately reflect the distribution of the fluorophore molecules then any measurement is compromised and a measurement based on two images is doubly compromised. Sometimes a global measurement, like the average intensity over a ROI, can be extracted from noisy data and a poor image is redeemed by using a large number of pixels. However noise increases the apparent vari-

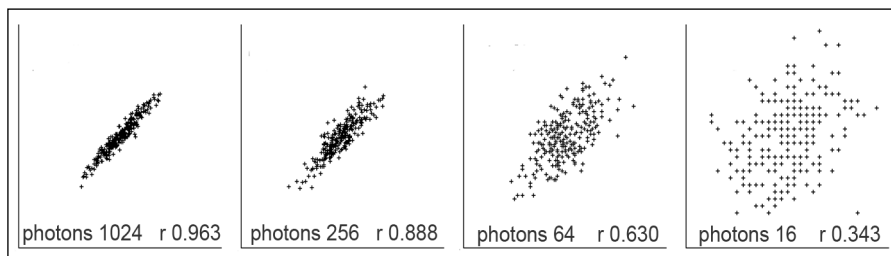


Figure 2:

Dependence of the measured colocalisation on noise. The scattergrams of simulated data illustrate how Poisson noise, measured by the number of quanta (photons), degrades the measured correlation of otherwise perfectly correlated data. As the number of photons falls and the Poisson noise increases, the dispersion of the data increases and the measured PCC drops.

ability in images, which will consequently reduce the measured PCC between two fluorophores (Figures 2 and 3) and the accuracy of the PCC is not restored even when the number of pixels is large. Simulations suggest that even with an average of 1024 quanta the PCC is still 4% below the true value (Figure 2). Simulations also show that the convergence of the measured PCC with the true value is asymptotic (Figure 3), making it difficult to acquire images of the quality needed to directly obtain a precise and accurate PCC. The need for high quality images has been noted before [4], but a remedy for extracting the true PCC from poor quality images is, to the best of our knowledge, new.

METHODS

Images were acquired using a Zeiss 510 Meta confocal microscope with an NA 1.4 objective using the linescan acquisition mode with the sequential use of two excitation wavelengths. The cells used were Jurkat T cells, clone 6.1, fixed in 4% formaldehyde.

All measurements and the creation of simu-

lated images used software based on a Semper6w kernel (Synoptics, UK). Further details are given in [2]. The scale bar uses a new format [5].

RESULTS

Since the failure of the PCC to accurately report the true colocalization is due to imperfections in the original images we can either try to obtain perfect images or use an estimate of image quality to correct the measured PCC [2]. Acquiring perfect images (see Table 1) involves detecting a large numbers of photons, although simulations show that the number required varies with the magnitude of the true intensity range [2], which is in practice hard to establish. Both the acquisition of perfect images and compensating for noise actually require some measure or estimate of image quality. Bizarrely, there is basically no provision for the measurement of image quality in image acquisition software. This means that the quality of an individual image depends upon the opinion, whim or bias of the microscope operator – often a poorly

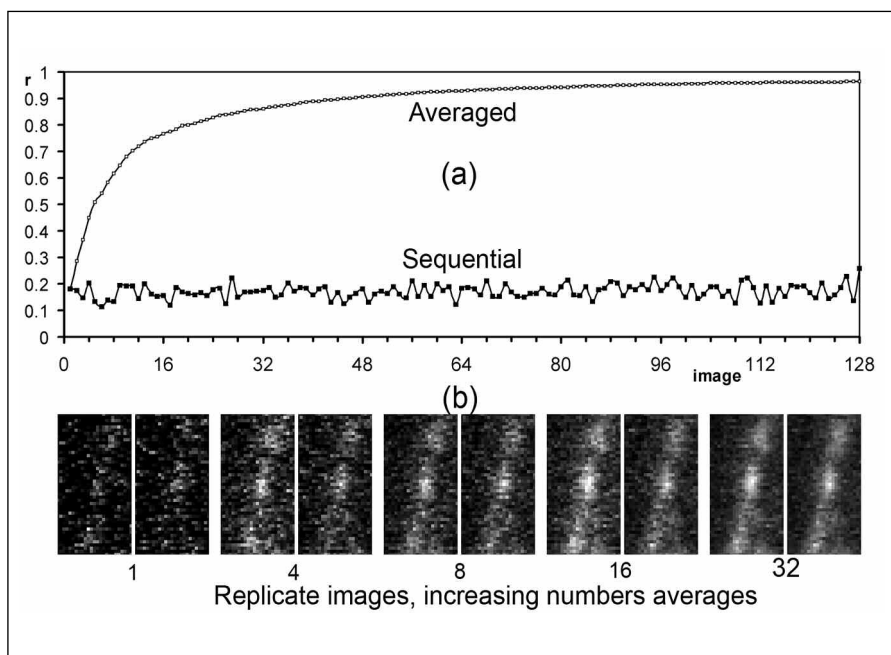


Figure 3:

Self colocalisation and image quality. The measured PCC measured between sequentially acquired images is consistent but wrong. Since the sequentially acquired images are of the same specimen the PCC should be 1. When images are averaged, the measured PCC progressively rises and approaches 1, a perfect match. The images used are shown (b) along with the number of images contributing to the average.

trained postgraduate student – which creates a very unstable basis for quantitative analysis [6].

An empirical method for estimating the quality of images in confocal line-scanning microscopy is to follow the changes to individual pixels as a line of pixels is rescanned. If the change is minimal then the images are good. We turn this approach into a practical quantitative measure by acquiring an image and a replicate (a second image taken immediately after the first) and comparing them using the PCC. Accordingly we call the method Replicate Based Noise Corrected Correlation (RBNCC). A comparison between replicate images has been used before in electron microscopy [7] and fluorescence microscopy [8] to assess image quality. Since the pair of images of each fluorophore are replicates their PCC would, in the absence of noise, be 1. Departures from 1, in a properly working system adhering to the conditions outlined in Table 1, are due to noise and the PCC between replicates is a good measure of image quality.

Using replicates for each fluorophore we obtain a self correlation for each fluorophore (r_{rr} and r_{gg}) and these are combined into a correction factor (C_{gr}):

$$C_{gr} = \frac{1}{\sqrt{r_{rr} \cdot r_{gg}}}$$

which is applied to the measured PCC (r_{mrg}) between the fluorophores producing a true or 'corrected' PCC (r_c) (Figure 4):

$$r_c = r_{mrg} \cdot C_{gr}$$

Overall the basis for our correction is that the accuracy with which two images of a fluorophore self-correlate also affects its subsequent correlation with a second fluorophore. In practice by simply acquiring a replicate image for each fluorophore it is possible to produce a corrected PCC that is almost identical to the PCC obtainable from noise-free images. Which, given that acquiring perfect images is rather both time and fluorophore consuming, represents a major advance of practical importance [2].

There are three additional features of the correction: (a) because we have two images for each fluorophore we can make four separate between fluorophore PCC measurements, the correction is applied to the mean of these four PCCs; (b) the same pixels are used for the between fluorophore PCC and the between replicate PCC; and (c) these pixels must have detectable levels of fluorescence.

PRACTICAL VALIDATION

Our correction for the PCC was derived empirically – we assumed the existence of a solution and tried various fiddles with test data. Originally we tried the simplest possible system: two copies of a set of the same data that were independently degraded by Poisson noise. *A priori* we knew that the true correlation must be 1 and could therefore test putative corrections – any plot that managed to combine the self correlations and the between dataset correlation and generate 1. Subsequently we increased the complexity of the original data,

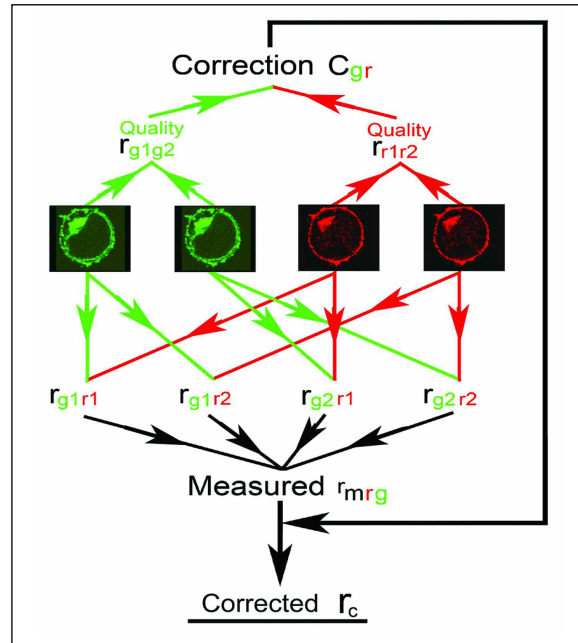


Figure 4: Schematic for making an accurate measurement of the PCC using RBNCC. Two images for each fluorophore are required and 6 measurements of the PCC are made. The replicate images are used to make a quality estimate (quality and quality) for each fluorophore and these are combined into a correction factor C_{gr} . The measured PCC between the different fluorophores (r_{mrg}) is obtained from the average of the PCC measurements made between all four possible combinations of the red and green images. The corrected PCC (r_c) is derived by combining the correction factor and the measured PCC.

by using Poisson noise based on different numbers of quanta (Figure 5) and with background noise with differing distributions. In practice RBNCC worked amazingly well and our simple correction compensated for both Poisson and background noise [2].

We also used confocal microscope images – here the noise present and the true PCC were unknown but we assumed that as the quality of the original images increased the measured PCC, without any correction, would tend towards the true PCC.

In practice the corrected PCC proved to be stable, varying little as the image quality changed and the measured PCC converged on the corrected PCC (Figure 6). The number of averages required to approach the corrected PCC is too large to be practical.

MATHEMATICAL VALIDATION

We later discovered that RBNCC was similar to a correction introduced by Spearman [9,10] developed for use with types of noise found in the social sciences.

In order to see if r_{mrg}/P , where $P = \sqrt{r_{rr} \cdot r_{gg}}$ is a legitimate formula, careful mathematical analysis is needed. In order to find out the average outcome of r as well as the average outcome of the product of self correlations r_{rr} , r_{gg} , so-called 'expected values' of the random errors need to be investigated. The errors in the denominators of each of the three mentioned correlations are easier to analyze if they are moved away from each denominator. This can be done by a high school Taylor series. After this, expected values can be calculated.

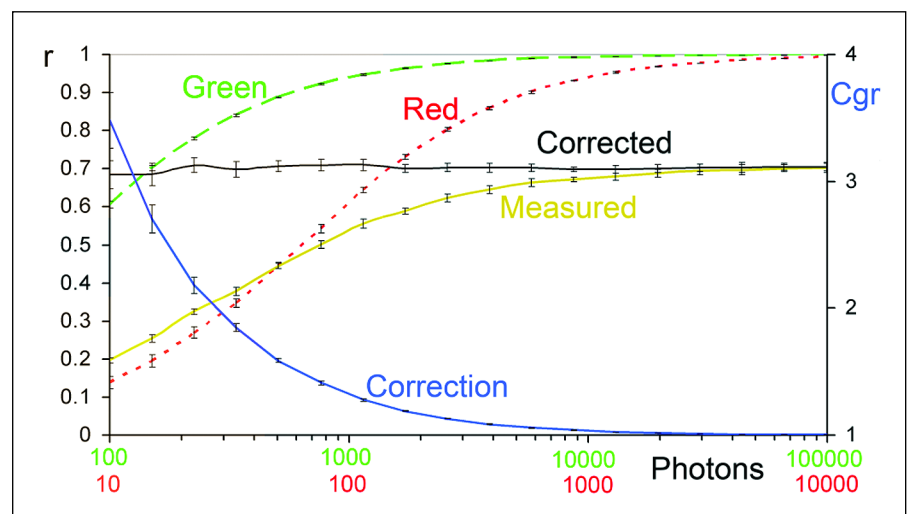


Figure 5: Testing RBNCC with simulated images. Pairs of red and green images with decreasing Poisson noise, produced by increasing the number of quanta (photons), show increasing self colocalization (red and green). As the quality of the images increases the calculated correction falls, from around 3.5 and finally approaches 1. The combination of the correction with the measured colocalization in the red and green images is used to calculate a corrected colocalization. Importantly the corrected colocalization is stable, despite huge changes in the number of quanta and therefore the Poisson noise. Note that the red images have 1/10 the number of photons used in the green images. Green, Red, Corrected and Measured are all correlations and use the left X axis. The secondary X axis on the right is used for the correction.

Such an analysis shows that errors in the denominator of r and in the denominator of r_{rr} times r_{gg} cause the same size of the percentage error. Hence, if (on average) $P = 0.5$, instead of being equal to 1, then (on average) the measured r needs to be doubled to obtain the true correlation. Provided errors in the red images are not correlated with errors in green images. But this result hinges on a simplifying assumption: namely that of neglecting errors in numerators of all involved correlations. However, by combining the Taylor series trick with a split-up of the correlations r_{mrg} , r_{rr} , r_{gg} into four terms that represent all the products possible between errors and signals, calculations are feasible. The result is: (a) the correction r/P is (on average) optimal; and (b) there are analytic formulas for the errors. The mathematical validation is new. The variance of r_{mrg}/P has been shown to be small for large datasets, like digital images [11].

DISCUSSION

All quantitative measurements are affected by the quality of the original data and despite its self evident importance in fluorescence microscopy image quality rests on the judgement of the microscope operator, typically a postgraduate student. Fluorescence microscopy is complex [3] and the selection of which cells to analyse and the quality of the images can undermine expensive, time-consuming and complex experiments.

We suggest that the comparison between replicate images [2] is an ideal basis for quality control and that scattergrams combined with the PCC ought to be made available at the time images are obtained, the only time deficiencies can be rectified. It is also important that images should have a quality fit for their scientific purpose, too low a quality generally undermines quantitation while excessive quality wastes resources.

We suspect that the measured PCCs generally understate the true PCC and even with images of apparently high quality the deficit averaged 23%, but with appreciable variability [2]. Accordingly we have considerable reservations about reported PCCs of nearly 1, 0.996 [12], since this would require that both images were perfect and that the true PCC was itself 1. It is however possible to achieve very high PCCs by judiciously selecting the pixels for analysis directly from a scattergram, rather than from a region of interest in the specimen. Since image quality affects the measured PCC the apparent objectivity provided by a quantitative measurement may be undermined by variations in image quality. Problems with image quality are unlikely to be limited to the measurement of colocalization.

In the context of colocalization measurements we have shown that it is possible to obtain accurate PCC from poor data, using RBNCC. Since poor data can be obtained more readily than perfect data and the photons required for a perfect image are sufficient for many poor images, RBNCC is ideal for live cell imaging, where the photobleaching and phototoxicity limit the number of available photons, or where short exposure times limit the

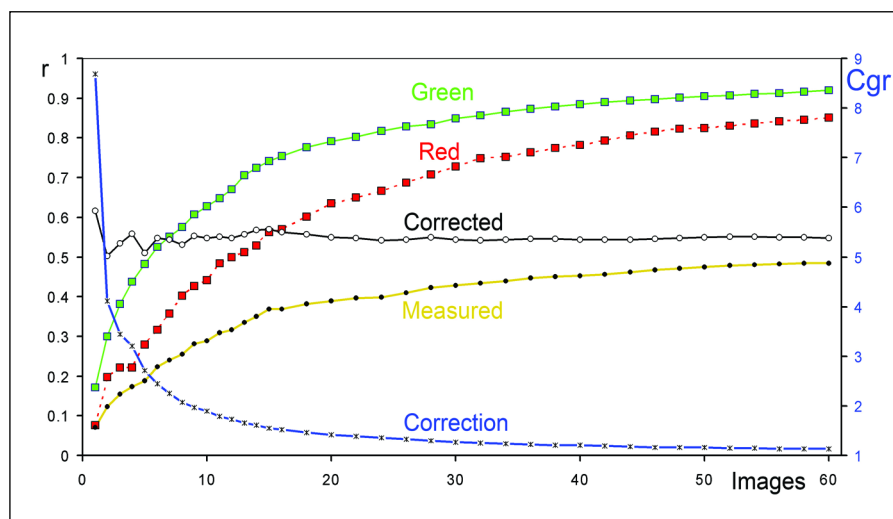


Figure 6:

RBNCC and confocal images. Sequentially acquired confocal images were used to generate images of increasing quality and a correction was calculated from the self colocalization measurements (red and green). The correction was applied to the measured colocalization and produced the corrected colocalization. Note that the corrected colocalization was consistent, even when the correction was high as 9. Green, Red, Corrected and Measured are all correlations and use the left X axis. The secondary X axis on the right is used for the correction.

Microscope	Images aligned in X, Y and Z Stability during image acquisition – no drift Same resolution
Data Quality	Data integrity – each image based on photons from a single fluorophore – no bleed through or autofluorescence Data falls within the detection range Free from Poisson noise Free from background noise Absence of FRET and quenching Stable, even illumination

Table 1:

Requirements for colocalization using the Pearson correlation coefficient.

number of obtainable photons.

While RBNCC corrects for noise it does not correct for misalignment of images and cannot compensate for data corruption, arising from crosstalk between the detector channels. Overall RBNCC is a highly efficient method for measuring the PCC accurately. It is not difficult to implement and the only practical requirement is the acquisition of four images, two pairs, instead of the usual two images. Since RBNCC does not require high quality images the acquisition of four images is actually a less onerous experimental requirement than the acquisition of the two high quality images otherwise required to accurately measure the PCC.

REFERENCES

- Manders, E. M. et al. Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *J. Cell Science* 103:857-862, 1992.
- Adler, J., Pagakis, S. N., Parmryd, I. Replicate Based Noise Corrected Correlation for Accurate Measurements of Colocalization. *J. Microscopy* 230(1):121-133, 2008.
- Pawley, J. B. *Handbook of biological confocal microscopy*. Springer, New York, 2006.
- Manders, E. M., Verbeek, F. J., Aten, J. A. Measurement of

co-localization of objects in dual-colour confocal images. *J. Microscopy* 169:375-382, 1993.

- Adler, J. The Unitary Scale Bar: human and machine readable. *J. Microscopy* 230(1):163-166, 2008.
- Pearson, H. The Good the Bad and the Ugly. *Nature* 447:138-140, 2007.
- Frank, J., Al-Ali, L. Signal-to-noise ratio of electron micrographs obtained by cross correlation. *Nature* 256:376-379, 1975.
- Costes, S.V. et al. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysics J.* 86:3993-4003, 2004.
- Spearman, C. The proof and measurement of association between two things. *Am. J. Psychology* 15:72-101, 1904.
- Spearman, C. Correlation calculated from faulty data. *Brit. J. Psychology* 3:271-295, 1910.
- Bergholm, F., Adler, J. and Parmryd, I. On the correction of correlation: a mathematical validation. Submitted.
- Zinchuk, V., Zinchuk, O., Okada T. Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: pushing pixels to explore biological phenomena. *Acta Histochem. Cytochem.* 40:101-111, 2007.

©2008 John Wiley & Sons, Ltd