Systems biology

Estimation of errors introduced by confocal imaging into the data on segmentation gene expression in *Drosophila*

Ekaterina Myasnikova¹,*; Svetlana Surkova¹; Lena Panok²; Maria Samsonova¹ and John Reinitz²

¹St Petersburg State Polytechnic University, St Petersburg, 195251, Russia and ²Stony Brook University, New York, 11794-3600, USA

Received on July 18, 2008; revised on November 25, 2008; accepted on November 27, 2008

Advance Access publication December 3, 2008

Associate Editor: Olga Troyanskaya

ABSTRACT

Motivation: Currently the confocal scanning microscopy of fluorescently tagged molecules is extensively employed to acquire quantitative data on gene expression at cellular resolution. Following this approach, we generated a large dataset on the expression of segmentation genes in the *Drosophila* blastoderm, that is widely used in systems biology studies. As data accuracy is of critical importance for the success of studies in this field, we took a shot to evaluate possible errors introduced in the data by acquisition and processing methods. This article deals with errors introduced by confocal microscope.

Results: In confocal imaging, the inevitable photon noise is commonly reduced by the averaging of multiple frames. The averaging may introduce errors into the data, if single frames are clipped by microscope hardware. A method based on censoring technique is used to estimate and correct this type of errors. Additional source of errors is the quantification of blurred images. To estimate and correct these errors, the Richardson–Lucy deconvolution method was modified to provide the higher accuracy of data read off from blurred images of the *Drosophila* blastoderm. We have found that the sizes of errors introduced by confocal imaging make up ∼5–7% of the mean intensity values and do not disguise the dynamic behavior and characteristic features of gene expression patterns. We also defined a range of microscope parameters for the acquisition of sufficiently accurate data.


Contact: myasnikova@spbcas.ru

Supplementary information: Supplementary data are available at Bioinformatics online.

1 INTRODUCTION

The success of systems biology studies largely depends on the availability of accurate and standardized quantitative datasets. Nowadays confocal scanning microscopy plays a vital role in systems biology offering a way for the acquisition of such datasets. In combination with fluorescent labeling this technique produces digital images of fixed or living tissues, from which quantitative data at single cell resolution can be extracted.

Although confocal scanning microscopy is a generally employed technique in biology, its use in systems biology studies presents a number of challenges. To quantify fluorescence, possible errors in instrument functionality, sample preparation and mathematical treatment of data have to be considered.

Recently, we acquired a large dataset on the expression of segmentation genes in the *Drosophila* blastoderm. This dataset characterizes the dynamics of formation of segmentation gene expression domains over the whole period of segment determination, that amounts to 1.5 h of development, has cellular resolution in space and 6.5 min resolution in time (Supplementary Material).

The data are stored in the FlyEx database (http://urchin.spbcas.ru/flyex/; http://flyex.ams.sunysb.edu/FlyEx/). Due to its high spatial and temporal resolution, the dataset is widely used by biologists to study the mechanism of pattern formation, infer regulatory interactions in the segmentation genetic network and develop new mathematical models (for references see http://urchin.spbcas.ru/flyex/refs.jsp). As the dataset accuracy is of critical importance for the success of these studies, we start to evaluate possible errors introduced in the data by acquisition and processing methods. In this article, we investigate the errors caused by confocal imaging.

Due to the discrete nature of photon detection and the presence of confocal pinhole, the Poisson distributed photon noise is among the major sources of errors in confocal imaging (Pawley, 2006; see Supplementary Material for more details). The common way to reduce the photon noise is the averaging of multiple separate scans. However, such an operation may also cause a loss of information at high and low intensities due to clipping of single scans. Image clipping is a form of signal distortion that is a consequence of a limited grayscale range of an image. For example, for 8-bit format the pixel values are cutoff once they exceed the 255 threshold. The dynamic range of confocal image is governed by gain and offset of the photomultiplier tube (PMT). The PMT gain linearly amplifies a weak signal, but exponentially increases the photon noise. The offset is always set negative and applied to subtract the background level of intensities. The adjustment of gain and offset defines the upper and lower thresholds of image clipping. The averaging of clipped images yields the biased means at high and low intensities; the errors take large values when the photomultiplier is adjusted to the limits of its sensitivity.

*To whom correspondence should be addressed.
The diffractive blurring of images is an additional source of data errors. Formally a blurred image is represented by a convolution of the original (latent) image with a point spread function (PSF), the response of an imaging system to a point object. According to optical theory, the size of PSF is governed by the numerical aperture of the optical system and the wavelength of light. However, experimental limitations, the lens alignment errors, light aberrations, electronic noise and the properties of the specimen itself, which vary widely, dictate that imaging cannot be performed at the theoretical maximum resolution of the microscope and the PSF should be defined in each particular experimental situation.

To recover the original image from the blurred one the operation called deconvolution is applied. Among currently available methods for image deconvolution the best restoration of an image distorted by Poisson distributed noise is provided by the maximum likelihood, or Richardson–Lucy (RL) algorithm (Dey et al., 2006; van Kempen and van Vliet, 2000). However, this method is very sensitive to errors in the PSF evaluation, and even given the precise PSF, does not provide the perfect and unique restoration of the original image, especially if the image objects have sharp edges. For a more detailed review of deconvolution algorithms see Supplementary Material.

Here, we characterize the contribution of errors introduced by confocal imaging to data on segmentation gene expression. We further describe how such experimental errors can be corrected and define a range of microscope parameters for the acquisition of sufficiently accurate data.

2 SYSTEM AND METHODS

To estimate the contribution of pixel saturation to data inaccuracy 12 wild-type (OregonR) Drosophila melanogaster blastoderm embryos were immunostained for the expression of one of four segmentation genes giant, even-skipped, hunchback and bicoid as described (Janssens et al., 2005; Kosman et al., 1997, 1998). Each embryo was scanned several times with the different combinations of gain and offset settings (Supplementary Table 1). Each embryo was additionally stained with an anti-histone H1-4 antibody (Chemicon) to mark the nuclei. Fluorescent labels used were Alexa Fluor 488 (gt and bcd), Alexa Fluor 555 (eve), Alexa Fluor 647 (hb) and Alexa Fluor 700 (histones) (Molecular Probes).

The images of 104 embryos from the FlyEx database were used to estimate errors caused by blurring and image segmentation. Each embryo was stained to detect the expression of three genes. Among them we considered 53 embryos stained for hb (Alexa Fluor 488), eve (555), 41 embryos for fushi-tarazu (ftz) (488), eve (555), sloppy-paired (slp) (647) and 10 embryos stained for bicoid (bcd) (647) and eve (488).

The quantitative gene expression levels in nuclei are obtained from the image with the use of a nuclear mask (Fig. 1d and e). The mask is a binary image in which all the pixels located within a nucleus are white and the rest pixels are black. The mask is superposed on the image and the values of pixels belonging to a nucleus are averaged.

2.1 Confocal microscopy

Embryo images were taken with the 20X Plan Apo dry objective ( numerical aperture 0.7) of a Leica TCS SP2 confocal system.

In the pixel saturation experiments the standard scanning procedure (Janssens et al., 2005; see also Supplementary Material) was modified. For each gene we image only one optical section of an embryo. This section was sequentially scanned 16 times and each scan was saved as a separate image (Fig. 1b). These images were averaged pixel-by-pixel and the resultant image was also saved (Fig. 1a and c). All 16 scans taken for one staining of a segmentation gene were pairwise compared to verify the absence of bleaching.

One and the same staining of a segmentation gene was scanned with different gain and offset settings to change the contrast of an image (Supplementary Table 1). The gain values ranged from 600 V to 920 V, offset was always negative and set from −4% to −45%. To enhance signal, the pinhole was opened to a diameter corresponding to 3 AU. All the other microscope parameters were fixed.

Images were sampled at a pixel size 0.4–0.48 µm, image size was on the average equal to 1200 × 500 pixels and a nucleus was composed of about 80 pixels.

2.2 Algorithm

2.2.1 Errors due to over- and under-saturation

The photon noise is the main source of noise in a single scan. This noise follows the Poisson distribution, that is characterized by variances proportional to the means. Figure 2b displays the relation between within-nucleus variance and nuclear mean in a single scan of a representative embryo. The within-nucleus variance is computed as the variance of all pixel values belonging to a nucleus. The inspection of the figure reveals that the growth of within-nucleus variance with the mean is only observed at the values of fluorescence intensity <150. The deviation from proportionality at higher intensity values may be referred to the well-known problem of image clipping. The images of segmentation gene expression are stored in an 8 bit format and have pixel intensity values ranging from 0 to 255. Thus if the value of any pixel in a single scan is beyond the [0 … 255] range, this value will be clipped to zero or 255, or under/over-saturated. The averaging of scans is performed pixel-by-pixel, therefore pixels clipped in some of the scans will not look saturated in the averaged image.

Besides pixel saturation, the residual nonlinearity of the microscope photon-counting circuitry is also responsible for the observed deviation from proportionality (Supplementary Material). However, the contribution of this effect is much less than that of the saturation. The nonlinearity of the photodetector response in the old confocal microscope models is a well-known fact (see for example Pawley, 1995), but this problem is virtually overcome in modern microscopes.

The formal statement of the problem is presented below.

(i) Over-saturation: the estimation of expression levels in spotted microarrays with over-saturated pixels was introduced in Ekstrom et al. (2004); Glasbey et al. (2007); Wit and McClure (2003). Taking a similar approach we say that the nonsaturated (latent) value of intensity in a single y, is censored above
at the threshold $c_b = 255$ if we observe

$$y_a = \begin{cases} 
  y & \text{if } y \leq c_a \\
  c_a & \text{if } y > c_a 
\end{cases}.$$ 

As it has been shown in Greene (2000), the averaging of censored data leads to inconsistent estimates of within-nucleus means and variances. In particular, at high intensities sample mean and variance underestimate the true values.

Let us assume that the nonsaturated fluorescence intensity in a nucleus, $\gamma$, is distributed normally with the mean $\mu$ and variance $\sigma^2$. The continuous normal distribution provides a good approximation to the discrete Poisson distribution at high values of the means. It is easy to show that mean and variance of $y_a$ are then given by

$$E(y_a) = -c_a \Phi(\alpha) + (1 - \Phi(\alpha))(\mu + \eta \sigma)$$

$$\text{Var}(y_a) = (1 - \Phi(\alpha))(1 - \delta + (1 + \delta)^2 \Phi(\alpha)) \sigma^2,$$

where $\alpha = (\mu - c_a)/\sigma$, $\eta = \phi(\alpha)/(1 - \Phi(\alpha))$, $\delta = \eta^2 - \eta \alpha$, $\phi$ and $\Phi$ are the density and cumulative distribution function of the standard normal distribution, respectively. The parameters $\mu$ and $\sigma$ are estimated for each nucleus using the method of moments to fit the measured within-nucleus means and variances to $E(y_a)$ and $\text{Var}(y_a)$.

So far we have ignored the effect of the nonlinearity of photon detection that leads to the skewness of the distribution of within-nucleus intensity values and hence makes the normal model (1) inadequate and the estimates of the distribution parameters inconsistent. It can be shown that the nonlinearity is well fit by the exponential saturation model (Pawley, 1995). To correct this effect for the observed intensity $y_a$ transform

$$y^* = B_0 - (B - B_0) \log \frac{B - y}{B - B_0} \quad \text{for } y > B_0,$$  

is applied, where $(B \geq 255)$ is the saturation asymptote, $(B_0 < B)$ is the lower saturation threshold.

To estimate parameters $B$ and $B_0$ we introduce the additional characteristics of the distribution, the fractions of clipped pixels per nucleus, defined as follows. Each pixel is an average of pixel values in $K$ scans, some of these values being cutoff at 0 or 255. Let $k^B_0$ and $k^B_{255}$ be a number of scans censored at 0 and 255, respectively, for $i$-th pixel belonging to $j$-th nucleus. Then the fractions of clipped pixels per nucleus are computed as $p^B_i = 1/(K_N) \sum_j k^B_{ij}$ and $p^B_{255} = 1/(K_N) \sum_j k^B_{255}$, where averaging is performed over $N_i$ pixels belonging to $j$-th nucleus. The parameters $B$ and $B_0$ are estimated by the least squares method fitting the probabilities $p^{255} = \Pr (\gamma \geq 255) = 1 - \Phi((255 - \mu)/\sigma)$ to these fractions. Given the current estimates of parameters $\mu_j$ and $\sigma_j$ of the intensity distribution in $j$-th nucleus, we minimize

$$\sum_j \left( p^{255}_j - p^B_{255,j} \right)^2,$$

with respect to $B$ and $B_0$. The values of these parameters are obviously the same for all the nuclei.

The observed intensity values transformed by (2) are used to find new estimates of the distribution parameters and the whole procedure is repeated until the sufficiently good fit of (3) is achieved.

(ii) Under-saturation: as a result of offset adjustment, a certain percent of intensities is subtracted from an image and any pixel value smaller than the substraction threshold is clipped and set to 0. This type of image distortion distorts the observed values of within-nucleus fluorescence intensity means. In this case, the observed intensity in a single scan is censored below the threshold $c_b$ and translated by its value

$$y_b = \begin{cases} 
  y - c_b & \text{if } y \geq c_b \\
  0 & \text{if } y < c_b 
\end{cases}.$$ 

The threshold $c_b$ is defined by the value of offset. The censoring at 0 can be corrected similarly to the censoring at high intensities, but the distribution of nonsaturated fluorescence intensities within a nucleus, $\gamma$, cannot be assumed normal any more. At low intensities we will consider the continuous Weibull and gamma distributions, which provide good approximation of the Poisson distribution (Wu, 2002). These distributions are close to exponential at low means and become very close to normal at high values of the means.

The Weibull distribution is described by the density function $f(x) = \rho/\lambda (x/\lambda)^{\rho-1} \exp(-x/\lambda)^{\rho}$, $x \geq 0$, with the mean $x^\Gamma(1 + 1/\rho)$ and variance $x^\Gamma(1 + 2/\rho) - x^\Gamma(1 + 1/\rho)$, where $\Gamma(a)$ is the gamma function. The mean and variance of censored random variable $y_b$ are then given by

$$E(y_b) = c_b \left( 1 - e^{-x/\lambda} \right) + x^\Gamma(1 + 1/\rho)$$

$$\text{Var}(y_b) = c_b^2 \left( 1 - e^{-x/\lambda} \right) + x^\Gamma(1 + 2/\rho) - x^\Gamma(1 + 1/\rho) - E^2(y_b),$$

where $y = (c_b/\lambda)^\rho$, $\Gamma(a, x)$ is the incomplete gamma function. The distribution parameters are estimated for each nucleus in the same way as the parameters of normal distribution in the case of over-saturation (1).

The gamma distribution is applied similarly and will not be considered here in detail.

2.2.2 Errors due to image blurring. The quantitative gene expression data on segmentation gene expression are the relative values computed by averaging the fluorescence intensity values over all the pixels composing a nucleus. Evidently, the precise detection of nuclear borders is an important condition for the acquisition of accurate quantitative data. However, even in the case of ideally precise mask the diffractive blurring distorts the intensity values at sharp borders and ignoring this fact leads to underestimation of the mean levels of fluorescence intensity. In this section we introduce a method to estimate data errors caused by image blurring.

Figure 3a illustrates the effect of blurring by considering a nucleus of an embryo stained for the expression of $hb$. In this figure, the values of fluorescence intensity of $Hb$ are the highest in the middle of the nucleus and decrease towards its edges and cytoplasm.

We believe that the observed bell-shaped distribution of intensity values can be explained for the most part by the diffractive blurring and partially by other optical and nonoptical effects. Indeed, the segmentation proteins are transcription factors and bind to many target genes (Ingham, 1988). A gene could occupy any radial position (Foe and Alberts, 1983; Wilkie et al., 1999) within a specific apical-basal plane of the Drosophila blastoderm nucleus.

Therefore the segmentation protein should show no preference to any part of the nucleus section recorded in the confocal image. The same should hold for the distribution of the fluorescence intensity values as antibody staining is known to be proportional to protein concentration. Nevertheless, we observe the bell-shaped distribution of intensities in a nucleus not only for $Hb$ but
Thus the PSF is completely defined by the value of pixel values. (which will be discussed below. borders detection, but, nevertheless, may contain some minor imperfections, image segmentation method, used for the acquisition of quantitative very sensitive to the accurate identification of intranuclear pixels. The for any other segmentation protein, suggesting that such a distribution has black diamonds. The pixels marked by crosses are not used in fitting.

for any other segmentation protein, suggesting that such a distribution has nonbiological nature and may be generated by the microscope optics.

Evidently, the extraction of quantitative data from such images is very sensitive to the accurate identification of intranuclear pixels. The image segmentation method, used for the acquisition of quantitative gene expression data and described in Janssens et al. (2005) (see also Supplementary Material), provides a very high precision of the nuclear borders detection, but, nevertheless, may contain some minor imperfections, which will be discussed below.

The blurring of an image can be described by the PSF, or the blur function, which is given for a confocal microscope by the first-order Bessel function of the first kind \( J_1 \) (Dey et al., 2006) as

\[
\text{PSF}(k, l) = \frac{2J_1(\alpha r(k, l))}{\alpha r(k, l)}
\]

where \( r(k, l) = \sqrt{k^2 + l^2} \) is the radial distance between two pixels in the image plane, \( \alpha \) is a blur parameter. Then a pixel value in the blurred image is equal to the convolution of the PSF with the pixel neighborhood

\[
u_{ij} = (u_{ij}^{\text{PSF}}) = \sum_{k,l} u_{ik,jl} \text{PSF}(k,l),
\]

where \( u_{ij} \) is the value of the same pixel in the original image. The parameter \( a \) defines the degree of blurring, the lower is \( a \) the more the image is blurred. Thus the PSF is completely defined by the value of \( a \).

To estimate the blurring at the nucleus boundary the original (unblurred) image is approximated by the step image, \( S \) (presented by dashed line in Fig. 3b and c). In this image all the pixels in the \( m \)-th nucleus \( \mathbb{N}_m \) have equal values, \( N_m \), and all the pixels in the cytoplasm \( C_m \) surrounding the nucleus take the same value \( C_m \). Given the blur parameter \( a \), parameters \( N_m \) and \( C_m \) are found by the least squares method that minimizes the sum of squared differences between the blurred step image \( S + \text{PSF} \) (solid line) and the observed image (black circles) for each nucleus separately

\[
R_{nm} = \sum_{i,j} (u_{ij}^{(S + \text{PSF})} - u_{ij})^2.
\]

Visual inspection of the images shows that segmentation gene products are not distributed uniformly within a nucleus, but are rather organized in randomly scattered patches, and hence the fluorescence intensity within a nucleus cannot be approximated by a constant value. Therefore the blur of the image is only estimated from the subset of within-nucleus pixels adjoining the boundary, \( \mathbb{N}_m \), united with the set of out-of-nucleus pixels, \( C_m \). The subset \( \mathbb{N}_m \) is composed of the pixels that are not more than 1 pixel away from the border. The blur parameter \( a \) is estimated by minimizing \( \sum_n R_{nm} \) with respect to \( a \) using the Powell optimization algorithm (Press et al., 1988). The summation is done over all the nuclei. At each step of the minimization parameters \( N_m \) and \( C_m \) are found anew via the solution of the linear minimization problem (7). Thus the set of parameters \( \{N_m,C_m\}_m \) and the parameter \( a \) are estimated in the same two-level optimization procedure.

Once the value of the parameter \( a \) has been found, the blurred image can be recovered by any standard deconvolution algorithm that requires a known PSF. We apply the iterative regularized RL algorithm (Dey et al., 2006), the basic idea of which is to calculate the most likely \( u_{ij}^{\text{PSF}} \) given the observed \( u_{ij} \) and known PSF. This is done by solving iteratively the equation

\[
u_{ij}^{(a+1)} = \frac{u_{ij}}{u_{ij}^{\text{PSF}}} \left( \frac{u_{ij}}{u_{ij}^{\text{PSF}}} - \psi \right) \cdot \frac{1}{1 - \rho \Delta u_{ij}^{(a+1)}},
\]

where the last factor comes from the Tikhonov–Miller (TM) regularization (van Kempen and van Vliet, 2000), the symbol \( \Delta \) stands for the Laplace operator \( \Delta u = u_{xx} + u_{yy} \) and \( \rho \) is a regularization parameter. The regularization is introduced to prevent the amplification of noise that occurs after several iterations of the nonregularized RL algorithm. It has been shown empirically that if the procedure (8) converges, it converges to the maximum likelihood solution (Shepp and Vardi, 1982).

It is well-known that the RL method incompletely removes the blur from the sharp object borders. As a consequence this method does not guarantee the exact restoration of a step-like image. To achieve more precise restoration the image is decomposed into step component and remaining signal. Thus it is sufficient to only apply the RL algorithm to the difference between the observed image \( u_{ij} \) and the blurred step image \( [(S + \text{PSF})] \) (black diamonds). If the observed image is well approximated by the blurred step one this difference does not have any more steep gradients and jumps of intensity at nuclear borders. As the RL method works only with positive pixel values, whereas the difference may take negative values, we add a positive constant, which will be subtracted after deconvolution. The recovered image (white circles) is then obtained as a sum of the step image and the recovered one this difference.

Fig. 3. (a) 2D distribution of pixel intensities over a nucleus and an island of surrounding cytoplasm in an image of an embryo stained for the expression of \( hb \). Black squares mark pixels attributed to the nucleus; white circles mark pixels imaging the cytoplasm area. The nucleus and cytoplasm areas are shown in the inset (b) and on the xy plane as a projection of corresponding pixel values. (c and d) Two sections of the graph (a) along AP- and DV-axes. Black squares mark pixels attributed to the nucleus and cytoplasm in the observed image; white circles mark the same pixels in the recovered image. Dashed step line is a section of the step image; solid line is a section of the 2D distribution of pixel intensities over a nucleus and an island of cytoplasm (or step image); dot line is a section of the observed image. (a) and (b) are the results of the exact restoration of a step-like image. To achieve more precise restoration the image is decomposed into step component and remaining signal. Thus it is sufficient to only apply the RL algorithm to the difference between the observed image \( u_{ij} \) and the blurred step image \( [(S + \text{PSF})] \) (black diamonds). If the observed image is well approximated by the blurred step one this difference does not have any more steep gradients and jumps of intensity at nuclear borders. As the RL method works only with positive pixel values, whereas the difference may take negative values, we add a positive constant, which will be subtracted after deconvolution. The recovered image (white circles) is then obtained as a sum of the step image and the recovered difference. This method, which will be further referred to as the step (or step RL) method, is illustrated in Figure 3b and c.

3 IMPLEMENTATION

3.1 Over- and under-saturation of averaged images

In Section 2.2.1, we introduce the method to correct mean intensity levels affected by over- and under-saturation. The method was tested on images distorted artificially by addition or subtraction of intensity units to imitate the saturation (Supplementary Material). The test confirmed the capacity of the method to reconstruct quantitative gene expression data with high accuracy. Of two models tested, the Weibull model proved to be adequate for correction of distortions at low intensities and is hereafter applied to real data.

Although the test demonstrated that quantitative gene expression data distorted by imitated over-saturation can be reconstructed with high accuracy, the nonlinearity of the PMT response is an
impediment to precise correction of quantitative data extracted from real images. The additional test described in Supplementary Material showed that the good adequacy of the model (1) can be only achieved in combination with the transformation (2) of the observed intensities in a single scan despite that the PMT nonlinearity is only apparent at the very highest intensities.

Considering that all the tests described above prove the high accuracy of the correction method, we apply it to estimate the errors introduced into real data by clipping. The quantitative data on $gt$ expression corrected to eliminate over- and under-saturation effect are shown in Figure 2a as an example. The means and variances were corrected for each scan and averaged over scans. After correction the dependance of within-nucleus variance on mean intensity is close to linear (Fig. 2b). This is typical for the Poisson distributed noise that prevails in a single scan confocal image. The absolute difference between corrected and saturated patterns both at high and low intensities does not exceed 12 intensity units (Fig. 2c).

It is also possible to estimate the range of microscope settings that provide the admissible level of errors in quantitative gene expression data. We calculate the absolute differences between measured and corrected mean fluorescence intensity levels for different values of microscope parameters. As it is evident from Figure 4, over-saturation becomes apparent at intensity levels about 150 and distortions in data become considerable for intensities >200. In such a situation the data errors may exceed 10 intensity units. The estimates of the absolute difference between corrected and saturated patterns both at high and low intensities does not exceed 12 intensity units (Fig. 2c).

In combination with the RL deconvolution and regularized filter provided the best restoration results with the lowest within-nucleus noise. We give preference to the RL method as the thorough analysis done in Dey et al. (2006) and van Kempen and van Vliet (2000) proved this method to be more effective in the presence of Poisson noise.

To estimate the inaccuracy of data quantification in real images, the step RL method was applied to confocal images of 104 embryos from the FlyEx database. The problem was solved in three phases. First, the blur parameter was estimated for each image, next the images were recovered to remove the blur and finally the data error caused by blurring was estimated as a difference in intensity levels between the recovered and FlyEx images.

The FlyEx images were previously segmented to detect the exact borders of nuclei and separate nuclei from cytoplasm (Janssens et al., 2005). As a result of segmentation all the pixels attributed to nuclei are explicitly listed. To apply the step method, we expand each nucleus area by a layer of two pixel width using the dilation operator (Gonzalez and Woods, 2002) as shown in Figure 3b. These extra pixels belong to cytoplasm and complement the list of nuclear pixels.

The blur parameter $a$ was estimated by minimizing the sum of $R_m$ given by (7) over all nuclei. The estimates fell within the limits of 1.35 and 1.6 $\mu$m$^{-1}$ for all the images. The estimates of the blur parameter from the images acquired in the Alexa Fluor 647 channel, 1.35–1.5 $\mu$m$^{-1}$, are somewhat smaller than those from images acquired with Alexa Fluor 555 and 488, which range between 1.45 and 1.6 $\mu$m$^{-1}$.

To cross-validate the results, we have experimentally measured the PSF from the images of sub-resolution fluorescent beads (Supplementary Material). The values of blur parameter in FlyEx images are little less (by 0.1–0.15 $\mu$m$^{-1}$) than those estimated from the bead images, as the latter do not capture out-of-focus light. Besides, the blur in FlyEx images is aggravated due to averaging of two optical sections, in which the boundaries of the nucleus may not exactly coincide.

The step image $S$ was determined by the set of parameters estimated in the same minimization procedure as the blur parameter $a$. Then the blurred step image ($S + PSF$) was subtracted from the observed one and the residual image was reconstructed using the regularized RL algorithm (8). The regularization parameter $\rho$ was taken equal to 0.02, as this is the parameter value that provides maximal noise reduction and at the same time prevents from negative pixel values.

To estimate the errors in quantitative gene expression data, we compute the difference between mean intensity levels in recovered and FlyEx images for each nucleus. These errors are positive (with the possible exception of the lowest intensities), as mean intensities read from the recovered image exceed those obtained from the blurred one, and increase linearly with intensity. For example, for the nucleus presented in Figure 3 the mean intensity computed from the observed pixel values is 113.6, while after restoration the within-nucleus mean increases up to 121.4. In comparison, the straightforward restoration by the RL method and blind deconvolution give similar estimates 116.3 and 117.1, respectively.

Figure 5a presents the sizes of errors in quantitative gene expression data extracted from three representative images of $hb$ expression pattern stored in FlyEx. The embryos used to obtain
As there is no ability to distinguish between segmentation errors and the main source of data inaccuracy is the segmentation procedure. Although the degree of blur depends on the wavelength and hence the necessary requirement for the accurate performance of the step method is to be very close to those obtained by the step RL method differing by at most one unit.

These problems are well recognized in literature and are demonstrated by the results of tests presented in Supplementary Tables 2 and 3. The level of noise in the tests is not higher than that observed in the FlyEx data and allows to achieve good correction of the data at the observed degree of blur.

One more factor that affects the error sizes is the adjustment of the blur parameter value. For example, if the parameter value is decreased from 1.6 µm⁻¹ to 1.5 µm⁻¹, the error estimates only increase by half intensity unit (Table 2 of Supplementary Material). In particular, this means that discrepancy of about 0.2 µm⁻¹ between the estimates of blur parameter by the step method and experimental one is not very essential. These two methods bring us to very similar estimates of data errors that differ by at most two units.

In order to additionally cross-validate the results, the errors were also computed using the step approach in combination with the regularized filtering. The error sizes estimated by this method proved to be very close to those obtained by the step RL method differing by at most 1 unit.

Obviously, the error sizes and accuracy of their estimation depend on the quality of a nuclear mask. If nuclear boundaries are detected not precisely enough, so that some extra cytoplasm pixels are assigned to a nucleus, the mean intensities will be additionally underestimated thereby increasing the error sizes. If detected nuclei are too small, the step method provides just minor corrections and the main source of data inaccuracy is the segmentation procedure. As there is no ability to distinguish between segmentation errors and errors introduces by blurring, the high quality of the nuclear mask is the necessary requirement for the accurate performance of the step method.

To confirm the high accuracy of masks constructed for the FlyEx dataset we have conducted a series of independent tests. These tests proved the low sensitivity of the masks to blurring and noise, as well as the good accuracy of boundary detection for those nuclei, which have sufficiently high mean fluorescence intensity values in at least one channel image (unpublished data). However, the accuracy of edge detection is limited by low sampling frequency, such that the diameter of a nucleus is 8–10 pixels. Besides, too high noise and too strong blur in images may also prevent the precise correction of the data. These problems are well recognized in literature and are demonstrated by the results of tests presented in Supplementary Tables 2 and 3. The level of noise in the tests is not higher than that observed in the FlyEx data and allows to achieve good correction of the data at the observed degree of blur.

One more factor that affects the error sizes is the adjustment of the blur parameter value. For example, if the parameter value is decreased from 1.6 µm⁻¹ to 1.5 µm⁻¹, the error estimates only increase by half intensity unit (Table 2 of Supplementary Material). In particular, this means that discrepancy of about 0.2 µm⁻¹ between the estimates of blur parameter by the step method and experimental one is not very essential. These two methods bring us to very similar estimates of data errors that differ by at most two units.

In order to additionally cross-validate the results, the errors were also computed using the step approach in combination with the regularized filtering. The error sizes estimated by this method proved to be very close to those obtained by the step RL method differing by at most 1 unit.

Obviously, the error sizes and accuracy of their estimation depend on the quality of a nuclear mask. If nuclear boundaries are detected not precisely enough, so that some extra cytoplasm pixels are assigned to a nucleus, the mean intensities will be additionally underestimated thereby increasing the error sizes. If detected nuclei are too small, the step method provides just minor corrections and the main source of data inaccuracy is the segmentation procedure. As there is no ability to distinguish between segmentation errors and errors introduces by blurring, the high quality of the nuclear mask is the necessary requirement for the accurate performance of the step method.

To confirm the high accuracy of masks constructed for the FlyEx dataset we have conducted a series of independent tests. These tests proved the low sensitivity of the masks to blurring and noise, as well as the good accuracy of boundary detection for those nuclei, which have sufficiently high mean fluorescence intensity values in at least one channel image (unpublished data). However, the accuracy of edge detection is limited by low sampling frequency, such that the diameter of a nucleus is 8–10 pixels. Besides, too high noise and too strong blur in images may also prevent the precise correction of the data. These problems are well recognized in literature and are demonstrated by the results of tests presented in Supplementary Tables 2 and 3. The level of noise in the tests is not higher than that observed in the FlyEx data and allows to achieve good correction of the data at the observed degree of blur.

One more factor that affects the error sizes is the adjustment of the blur parameter value. For example, if the parameter value is decreased from 1.6 µm⁻¹ to 1.5 µm⁻¹, the error estimates only increase by half intensity unit (Table 2 of Supplementary Material). In particular, this means that discrepancy of about 0.2 µm⁻¹ between the estimates of blur parameter by the step method and experimental one is not very essential. These two methods bring us to very similar estimates of data errors that differ by at most two units.

In order to additionally cross-validate the results, the errors were also computed using the step approach in combination with the regularized filtering. The error sizes estimated by this method proved to be very close to those obtained by the step RL method differing by at most 1 unit.

4 DISCUSSION

The standard way to acquire quantitative gene expression data at cellular resolution is to stain a specimen with fluorescent antibodies, scan it and process the resulting confocal images. Evidently all these procedures introduce errors in data. In this article, we have characterized and estimated the error in the quantitative data on segmentation gene expression introduced by confocal imaging. Due to high spatial and temporal resolution this data have recently received an increased amount of attention, and so it is important to scrutinize its accuracy.

The importance of the application of statistical methods to evaluate the quality of confocal images, as well as to estimate its dependence on microscope performance was previously recognized (Zucker and Price, 2001a,b). The main focus of these and other studies was the photon noise in images and the ways of its reduction. Here we consider other types of errors introduced by confocal microscope, namely errors caused by clipping of single scans and blurring.

The common way to reduce the photon noise is averaging of a number of independent confocal scans. Although the averaging of scans improves the signal-to-noise ratio of an image, it may cause image distortions, if single scans are clipped by microscope software. The quantitative gene expression data read from such an averaged image will be biased at high and low intensities. We have developed a method based on a censoring technique to estimate and correct errors in the quantitative data caused by over- and under-saturation. The method allows us also to estimate the acceptable range of microscope parameters that ensure the good accuracy of the data. It has been shown that there should be a reasonable balance between attempts to increase the dynamic range of an image and the loss of data accuracy at high gain and absolute values of offset.

To extract the quantitative data on segmentation gene expression from confocal images a binary nuclear mask was applied. The mask outlines the nucleus borders and isolates the areas from which the data are read. Even if the mask is highly accurate, the blurring of an image due to diffraction scattering distorts the intensity values at sharp borders. This results in underestimation of the mean levels of fluorescence intensity and hence gives rise to inaccuracy in the quantitative data. Evidently, even very small errors in the detection of nuclear borders may lead to considerable errors in mean intensities. Precise restoration of the image raises two main problems: estimation of the blurring PSF and elimination of data inaccuracy at high gain and absolute values of offset.
of distortions at the sharp borders of nuclei. The nearly random distribution of segmentation proteins gave us an opportunity to evaluate the errors introduced by blurring more accurately than the direct deconvolution of the observed blurred image. Our method extracts a component that has no abrupt jumps of intensity and the estimated PSF is only applied for the deconvolution of this component. As a deconvolution method, we apply the RL algorithm that proved to provide the best restoration of images distorted by Poisson distributed noise (Dey et al., 2006; van Kempen and van Vliet, 2000).

We have found that confocal scanning indeed introduces errors into the data, but the sizes of these errors are relatively small. For example, at intensities of about 200 the size of error caused by blurring is equal to 12 units at maximum, i.e. makes up ~6% of the mean intensity value (Fig. 5). Pixel saturation also introduces errors, which do not exceed 10–15 intensity units (5–7% of mean intensity) at reasonable values of the PMT parameters (Fig. 4).

It is necessary to make sure that errors in the data, as small as they might be, do not disguise the dynamic behavior and characteristic features of the pattern. Fortunately, in our dataset confocal scanning introduced quite minor distortions in the patterns of segmentation gene expression. This can be demonstrated by considering the Bcd gradient, which, due to its shallow profile, is especially sensitive to intensity distortions. As is evident from inspection of Figure 5, the correction of errors introduced by blurring shifts the position, where the Bcd concentration is 50% of maximum, to less than half a nucleus anteriorly.

Clearly, the proper choice of microscope parameters accounts for the small sizes of errors in the data. In the FlyEx dataset, the images were acquired with gain not exceeding 800 V and offset kept within interval from −5% to −25%. These are normally used parameter ranges and according to our results they guarantee good accuracy of data.

As was mentioned above, the confocal microscope is not the only source of error in the acquisition of quantitative gene expression data at cellular resolution. Among other sources of error are background staining, non-accurate embryo orientation, image and data processing errors, just to mention a few. The final goal of our study is to estimate all these errors and discriminate the natural biological variability from experimental noise originating from all possible sources.

In addition, it is worth mentioning that the correction methods developed in this work can be easily adapted to a wide range of biological images obtained with confocal microscopy. The general-purpose method for the correction of pixel saturation requires all the confocal scans to be saved without averaging. Our modification of deconvolution algorithm is applicable to other images displaying gene expression in nucleus or cell, which usually look like objects with sharp edges. The only additional requirement is the distribution of gene products, that should not show any preference to any part of these objects. Besides the image quality has to be good enough to construct a high-precision nuclear mask. For example, we have successfully applied the step method to correct the data extracted from the FlyEx images obtained in histone channel (Supplementary Material).

ACKNOWLEDGEMENTS

We thank A. Samsonov and G. Stein for the helpful comments.

Funding: National Institutes of Health (RR07801); Civilian Research & Development Foundation (Grant Assistant Program RUB1-1578); Netherlands Organization for Scientific Research and Russian Foundation for Basic Research (project 047.011.2004.013); Russian Foundation for Basic Research (grants 08-01-00315a, 08-04-00712a).

Conflict of Interest: none declared.

REFERENCES


