A quantitative comparison of image restoration methods for confocal microscopy[1]

G. M. P. VAN KEMPEN,* L. J. VAN VLIET,* P. J. VERVEER† & H. T. M. VAN DER VOORT‡

Pattern Recognition Group, Faculty of Applied Physics, Delft University of Technology, Lorentzweg 1,

2628 CJ Delft, The Netherlands

†*Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11,*

D-37077 Göttingen, Germany

‡*Scientific Volume Imaging B.V., J. Geradtsweg 181, 1222 PS Hilversum, The Netherlands*

Summary

In this paper, we compare the performance of three iterative methods for image restoration: the Richardson–Lucy algorithm, the iterative constrained Tikhonov–Miller algorithm (ICTM) and the Carrington algorithm. Both the ICTM and the Carrington algorithm are based on an additive Gaussian noise model, but differ in the way they incorporate the non-negativity constraint. Under low light-level conditions, this additive (Gaussian) noise model is a poor description of the actual photon-limited image recording, compared with that of a Poisson process. The Richardson–Lucy algorithm is a maximum likelihood estimator for the intensity of a Poisson process. We have studied various methods for determining the regularization parameter of the ICTM and the Carrington algorithm and propose a (Gaussian) prefiltering to reduce the noise sensitivity of the Richardson–Lucy algorithm. The results of these algorithms are compared on spheres convolved with a point spread function and distorted by Poisson noise. Our simulations show that the Richardson–Lucy algorithm, with Gaussian prefiltering, produces the best result in most of the tests. Finally, we show an example of how restoration methods can improve quantitative analysis: the total amount of fluorescence inside a closed object is measured in the vicinity of another object before and after restoration.

1. Introduction

The analysis of the three-dimensional (3-D) structure of tissue, cells and cellular constituents plays a major role in biomedical research. Three-dimensional images, often acquired with confocal fluorescence microscopes, are vital

to this analysis. However, the imaging properties of a confocal microscope give rise to a blurring phenomenon similar to the one in a conventional microscope, but with a reduced range (van Resandt *et al.*, 1985). The resulting distortions hamper subsequent quantitative analysis. Therefore, operations that invert the distortions of the microscope improve this analysis.

This inversion in the presence of noise is known to be a very difficult problem. In fact, the restoration of information that is severely suppressed by this blurring is known to be an ill-posed problem (Tikhonov & Arsenin, 1977). Therefore, *a priori* knowledge about the noise and object will regularize the restoration.

In previous work (van Kempen *et al.*, 1996) the performance of the iterative constrained Tikhonov–Miller (ICTM) algorithm to restore diffraction-induced distortions was compared with the Richardson–Lucy algorithm. Quantitative texture measurements, based on the grey-value distance transform, showed that the results improved when applied to images after restoration. The results showed that the performance of the Richardson–Lucy algorithm, measured by the mean-square-error and I-divergence criterion, was in most circumstances better than the ICTM result.

The use of the ICTM restoration method was motivated by the linear system model used to describe the imaging properties of a confocal microscope. In this model, the image is a convolution of the object function with the point spread function of the microscope and distorted by additive Gaussian noise. This image formation model breaks down on images with a low signal-to-noise ratio, where the additive noise model is a poor description of the actual photon-limited image recording. Under these circumstances, the noise characteristics are better described by a Poisson process, which motivates the use of restoration methods optimized for Poisson noise distorted images.

In this paper, we compare the Richardson–Lucy algorithm, which is a maximum likelihood estimator for the intensity of a Poisson process prior to distortion (Shepp & Vardi, 1982), with the ICTM and the Carrington algorithm (Carrington, 1990). Furthermore, we have tested different methods for determining the regularization parameter of both the ICTM and the Carrington algorithm. Previous work (van Kempen *et al.*, 1996) showed that a regularization parameter inversely proportional to the signal-to-noise ratio causes too much smoothing. Furthermore, we examine a method to reduce the noise sensitivity that is frequently observed with the Richardson–Lucy algorithm (Snyder & Miller, 1991).

Finally, we show an example of how restoration methods can improve quantitative analysis: the total amount of fluorescence inside a closed object is measured in the vicinity of another object before and after restoration.

A vast amount of literature is published on the restoration of microscope images, using restoration methods such as the Richardson–Lucy (Snyder *et al.*, 1993; Holmes *et al.*, 1995), ICTM (van der Voort & Strasters, 1995), the Carrington algorithm (Carrington, 1990; Carrington *et al.*, 1995) and maximum *a posteriori* algorithms (Joshi & Miller, 1993). Whereas these methods have just come within reach in terms of computational complexity, they have been shown to significantly improve the (quantitative) analysis of microscope images (van der Voort & Strasters, 1995). Owing to their nonlinear nature and application of prior knowledge, these methods are capable of partially restoring data from missing frequencies, as induced by the 'missing cone' of the 3-D OTF of incoherent light microscopes (Carrington, 1990; Shaw & Rawlins, 1991; Holmes *et al.*, 1995). Therefore, they significantly reduce the diffraction-induced distortion found in confocal and conventional 3-D images (Shaw & Rawlins, 1991; Holmes *et al.*, 1995).

2. Image restoration

2.1. Image acquisition

The incoherent nature of emitted fluorescence light allows us to model the image formation of the confocal fluorescence microscope as a convolution of the object function $f(x)$ with the point spread function (PSF) of the microscope $h(x)$, with x being a 3-D coordinate in the space X. The image $g(x)$ formed by an ideal noise-free confocal fluorescence microscope can thus be written as

$$g(x) = \int_X h(x - \chi)f(\chi)\,\mathrm{d}\chi. \tag{1}$$

Owing to the photon nature of light and its effect on $f(x)$, $g(x)$ is distorted by noise. Noise, caused by photon counting (Poisson noise), by the readout of the detector (Gaussian)

and by the analog-to-digital conversion (uniform), disturbs the image (Art, 1995). The performance of state-of-the-art detectors is limited by photon counting noise (Art, 1995). Furthermore, it is common in fluorescence microscopy to measure a nonzero background level, arising from auto-fluorescence, inadequate removal of fluorescent staining material, offset levels of the detector gain or other electronic sources. We model the noise distortion and background here in a general form:

$$m(x) = N(g(x) + b(x)) \tag{2}$$

with $m(x)$ being the recorded image, $b(x)$ the background and $N(\cdot)$ the noise distortion function.

2.2. Restoration methods

Generally, restoration methods yield an estimate of the original image $\hat{f}(x)$ given an imaging model, a noise model and additional criteria. These criteria depend on the imposed regularization and constraints implied on the solution found by the restoration algorithm. Although the methods we investigate in this paper share the imaging model they differ significantly owing to the different modelling of noise distortion on the image, imposed constraints and regularization.

The Richardson–Lucy algorithm computes the maximum likelihood estimator of the intensity of a translated Poisson process (Snyder & Miller, 1991). In case of additive Gaussian noise, the maximum likelihood criterion results in minimizing a mean-square-error criterion. Both the ICTM and the Carrington algorithm are constrained, regularized mean-square-error restoration methods for finding a non-negative solution for images distorted by additive noise. However, they differ in the way the non-negativity constraint is incorporated in the algorithms.

Poisson noise: the Richardson–Lucy algorithm. A confocal microscope acquires an image of an object by scanning the object in three dimensions. At each point of the image, the emitted fluorescence light from the object is focused on the detector. (This light is converted by a photomultiplier tube (PMT) into an electrical signal, and represented by a discrete value after an A/D conversion (Pawley, 1990).) Under low light-level conditions, the PMT detector behaves essentially as a photon counter. This conversion of fluorescence intensity to a discrete number of detected photons is described statistically as a Poisson process. Under these conditions, the image formation of a confocal microscope is described as a translated Poisson process (Snyder & Miller, 1991). A translated Poisson process models the measured data acquired from an underlying, unobservable Poisson process when the measurements are imperfect and in the form of a Poisson process. This process transforms a Poisson $F(x)$ with an intensity function $f(x)$ (representing the 'true'

object) into a Poisson process $m(x)$ subject to a conditional probability density function, which is in our case equal to the PSF $h(x)$. The log likelihood function of such a Poisson process is given by (Snyder & Miller, 1991)

$$L(f) = -\int_X g(x)\,dx + \int_X \ln[g(x) + b(x)]m(x)\,dx \quad (3)$$

where we have dropped all terms that are not dependent on $f(x)$ and with

$$g(x) = \int_X h(x - \chi)f(\chi)\,d\chi. \quad (4)$$

The conditional expectation of a Poisson process $F(x)$ given the measured data $m(x)$ is defined as

$$E[F(x)|m] = \int_X \left[\frac{h(\chi - x)}{g(\chi) + b(\chi)}\right] m(\chi)\,d\chi. \quad (5)$$

The maximum likelihood estimator (MLE) $\hat{f}(x)$ for restoring $f(x)$ given $h(x)$ and $m(x)$ can be found using the EM algorithm, as described by Dempster *et al.* (1977). This iterative algorithm was used first by Vardi *et al.* (1985) for computing the MLE of the intensity of a Poisson process. The EM algorithm requires two steps for each iteration, an expectation (E) step and a maximization step (M). In the expectation step, the conditional expectation $Q(\cdot)$ of the log likelihood function of $F(x)$ is determined:

$$L_X(f) = -\int_X f(x)\,dx + \int_X \ln[F(x)]f(x)\,dx$$

given the measured data $m(x)$ and the kth estimate $\hat{f}_k(x)$ of $f(x)$,

$$Q(f(x)|\hat{f}_k(x)) = \int_X f(x)\,dx + \int_X E[F(x)|m(x),\hat{f}(x)]\ln(f(x))\,dx. \quad (6)$$

Using (5), we obtain

$$Q(f(x)|\hat{f}_k(x)) =$$
$$\int_X f(x)\,dx + \int_X \left(\int_X \left[\frac{h(\chi - x)}{g(\chi) + b(\chi)}\right]m(\chi)\,d\chi\right)\ln(f(x))\,dx. \quad (7)$$

The M-step maximizes this function $Q(\cdot)$ over the set of admissible $f(x)$. Using $\partial Q/\partial f(x) = 0$ for $f = \hat{f}_{k+1}$ yields (Snyder *et al.*, 1993)

$$\hat{f}_{k+1}(x) = \hat{f}_k(x) \int_X \left[\frac{h(\chi - x)}{\int_X h(\chi - x')\hat{f}_k(x')\,dx' + b(\chi)}\right] m(\chi)\,d\chi. \quad (8)$$

The EM algorithm ensures a non-negative solution when a non-negative initial guess $\hat{f}_0(x)$ is used. Furthermore, the likelihood of each iteration of the EM algorithm will strictly increase to a global maximum (Snyder & Miller, 1991). The EM algorithm (8) is identical to the Richardson–Lucy algorithm as derived by Richardson (1972). Although accelerated and regularized versions of the Richardson–Lucy algorithm exist (Holmes & Liu, 1991; Joshi & Miller,

1993), none of these algorithms incorporate the background term $b(x)$.

From previous experiments (van Kempen *et al.*, 1996), we observed that the Richardson–Lucy algorithm is very sensitive to noise. We use a method that reduces this noise sensitivity by suppressing those parts of the image spectrum that do not contain any signal information (or where the noise contribution is far larger than the signal contribution). These frequencies will prevent signal recovery and only amplify noise in the final result. These (high-frequency) parts of the spectrum can be suppressed by convolving the recorded image with a Gaussian. We have compensated for this smoothing of this convolution in the Richardson–Lucy algorithm by convolving the PSF as well with the same Gaussian.

Additive noise distortion: the ICTM and the Carrington algorithm. Both the ICTM and the Carrington algorithms are based on the assumption that the general noise distortion function $N(\cdot)$ can be modelled as an additive noise function (Carrington, 1990, Lagendijk & Biemond, 1991):

$$m'(x) = m(x) - b(x) = \int_X h(x - \chi)f(\chi)\,d\chi + n \quad (9)$$

with n the additive noise distortion with zero mean. For images with a relatively high signal-to-noise ratio, the additive noise model can be motivated by the central limit theorem (Snyder & Miller, 1991): under these circumstances, the distribution of a Poisson process can be approximated by a space variant Gaussian distribution.

Finding an estimate $\hat{f}(x)$ from (9) is known as an ill-posed problem (Tikhonov & Arsenin, 1977). To solve this ill-posedness, Tikhonov defined the regularized solution $\hat{f}(x)$ of (9) as the one that minimizes the well-known Tikhonov functional (Tikhonov & Arsenin, 1977):

$$\Phi(f) = \left\|m'(x) - \int_X h(x - \chi)f(\chi)\,d\chi\right\|^2$$
$$+ \lambda\left\|\int_X r(x - \chi)f(\chi)\,d\chi\right\|^2 \quad (10)$$

with λ the regularization parameter and

$$\|f(x)\|^2 = \int_X |f(x)|^2\,dx.$$

The function $r(x)$ is known as the regularization filter. The Tikhonov functional consists of a mean-square-error criterion and a stabilizing energy bound. The functional suppresses solutions of $\hat{f}(x)$ that oscillate wildly due to spectral components outside the bandwidth of h. Minimization of (10) yields the well-known Tikhonov–Miller (TM) solution $a(x)$,

$$\hat{F}(\omega) = A(\omega)M'(\omega) \quad \text{with} \quad A(x) = \frac{H^*(\omega)}{\|H(\omega)\|^2 + \lambda\|R(\omega)\|^2} \quad (11)$$

with * denoting the conjugate operator, and all capitals the Fourier transform of the corresponding function. Although this solution requires modest computational efforts, it is very sensitive to errors in the estimation of the PSF, causing ringing artefacts (van der Voort & Strasters, 1995). Furthermore, the solution may contain negative values. This property of the TM inversion is a major drawback, since the intensity of an imaged object represents signal energy which is positive. We have compared two iterative algorithms, the ICTM and the Carrington algorithm, that incorporate this non-negativity constraint in solving (10).

The ICTM algorithm. The ICTM algorithm finds the minimum of (10) by using conjugate gradients. The so-called conjugate gradient direction of (10) is given by

$$p_k(x) = r_k(x) + \gamma_k p_{k-1}(x), \qquad \gamma_k = \frac{\|r_k(x)\|^2}{\|r_{k-1}(x)\|^2} \qquad (12)$$

with $r_k(x)$ denoting the steepest descent direction,

$$r_k(x) = -\tfrac{1}{2}\nabla_{\hat{f}}\Phi(\hat{f}(x)) = \int_X h(\chi - x)(m'(\chi) - g(\chi))\,d\chi - \lambda\hat{f}(x)$$

with $g(x)$ defined by (4). A new conjugate gradient estimate is now found with

$$\hat{f}_{k+1}(x) = \hat{f}(x) + \beta_k p_k(x). \qquad (13)$$

In the absence of a non-negativity constraint, the optimal step size β_k can be calculated analytically. However, in the presence of such a constraint, the optimal β_k must be searched for iteratively. In our implementation, a golden section rule line-search (Press *et al.*, 1992) is employed to find β_k. The ICTM algorithm, therefore, consists of a main iterative loop in which the conjugate directions are computed, and a subiterative loop in which β_k is optimized. The latter requires one blurring operation per subiteration.

The Carrington algorithm. Given the Tikhonov functional (10) (without the regularization operator) and its derivative,

$$\tfrac{1}{2}\nabla_f\Phi(f(x)) = \int_X h(\chi - x)(g(\chi) - m'(\chi))\,d\chi + \lambda f(x),$$

we want to find the non-negative value of $\hat{f}(x)$ that minimizes this functional. Then the Kuhn–Tucker conditions should apply (Carrington, 1990):

$$\nabla_f\Phi = 0 \text{ and } f > 0 \quad \text{or} \quad \nabla_f\Phi \geq 0 \text{ and } f = 0. \qquad (14)$$

Then from (14) the Kuhn–Tucker conditions follow:

$$\hat{f}(x) = \int_X h(\chi - x)c(\chi)\,d\chi \quad \text{if} \quad \int_X h(\chi - x)c(\chi)\,d\chi > 0 \qquad (15)$$

and

$$\hat{f}(x) = 0 \quad \text{if} \quad \int_X h(\chi - x)c(\chi)\,d\chi \leq 0 \qquad (16)$$

where

$$c(x) = \frac{1}{\lambda}\left(m'(x) - \int_X h(x - \chi)f(\chi)\,d\chi\right) = \frac{1}{\lambda}(m'(x) - g(x)).$$

From (15) and (16) now follows

$$\hat{f}(x) = \max\left(0, \int_X h(\chi - x)c(\chi)\,d\chi\right). \qquad (17)$$

On the domain X_+ where the values $\int_X h(\chi - x)c(\chi)\,d\chi > 0$ we then obtain, after insertion into (10), that

$$\int_X h(x - \chi)\left\{\int_{X_+} h(x' - \chi)c(x')\,dx'\right\}d\chi - m'(x) + \lambda c(x) = 0 \qquad (18)$$

where we have assumed that $h(\cdot)$ is non-negative. This is equivalent to minimizing the functional $\Psi(\cdot)$,

$$\Psi(c) = \tfrac{1}{2}\left\|\int_{X_+} h(\chi - x)c(\chi)\,d\chi\right\|^2 - \int_X c(x)m'(x)\,dx + \tfrac{1}{2}\lambda\|c(x)\|^2. \qquad (19)$$

$\Psi(c)$ is strictly convex and twice continuously differentiable. A conjugate gradient algorithm, similar to (12) and (13), can be used to minimize (19). In this case β_k can be found using Newton's rule. No blurring operations are necessary for the subiterations, resulting in a much faster line-search than in the case of ICTM.

2.3. Performance measures

In previous work (van Kempen *et al.*, 1996) we have used the mean-square-error (*MSE*) and the I-divergence as performance measures. The *MSE* is given as

$$MSE(f, \hat{f}) = \int_X |f(x) - \hat{f}(x)|^2\,dx. \qquad (20)$$

The mean-square-error measures the difference in energy between the two compared signals. Csiszár (1991) has introduced the I-divergence,

$$I(a, b) = a\ln\left[\frac{a}{b}\right] - (a - b),$$

to measure the distance of a function b to a function a. He has postulated a set of axioms of regularity (consistency, distinctness and continuity) and locality that a distance measure should possess. He concluded that for functions that are required to be non-negative, the I-divergence is the only consistent distance measure. For real-valued functions, having both negative and positive values, the *MSE* is the only consistent choice. Snyder *et al.* (1992) have shown that maximizing the mean of the log likelihood of (3) is equal to

minimizing Csiszár's I-divergence,

$$I(f,\hat{f}) = L(f) - E[L(\hat{f})] = \int_X \left(\ln\left[\frac{g(x)}{\hat{g}(x)}\right] g(x) - g(x) + \hat{g}(x) \right) dx.$$

(21)

Although the investigated image restoration algorithms minimize the *MSE* or I-divergence measure, these measures are not conclusive when the goal of image restoration is to improve the quantitative analysis of microscopic images. Therefore, we also examined the performance of these measured with an analysis-based performance measure.

The total amount of fluorescence inside a closed object is in many applications a useful measurement. It could for example represent the total amount of DNA inside a cell nucleus. Image blur, however, does not permit this measurement when objects lie close together. In this situation, image restoration algorithms can be applied, prior to the measurement, to reduce this blur. A useful performance measure is how the intensities are being reshuffled without mixing intensities that came from different objects. We have implemented this performance measure by assigning the intensity of a pixel to the object with the shortest distance from the pixel to the object's surface.

3. Experiments and results

We have performed two simulation experiments. The first experiment tests the performance of the Richardson–Lucy, ICTM and Carrington algorithms on restoring a sphere that is convolved with a confocal PSF and distorted with Poisson noise. The performance of the algorithms is measured with the *MSE* with I-divergence as a function of the signal-to-noise ratio, which is defined in this section. The second experiment tests the ability of the three algorithms to improve the quantitative measurement of the total amount of fluorescence inside one sphere in the vicinity of a second sphere.

3.1. Object and noise generation

We generated the spheres using an analytical description of their Fourier transform, as given by van Vliet (1993) in spherical coordinates u, v, w,

$$S_{\text{sphere}}\left(\frac{u}{r}, v, w\right) = \frac{-6\pi u \cos(2\pi u) + 3\sin(2\pi u)}{8\pi^3 u^3}$$

(22)

with r the radius of the sphere. The Fourier transform of the sphere is multiplied by a Gaussian transfer function (sigma of 1 pixel in the spatial domain) to ensure bandlimitation. Generated in this way, the spheres are free from aliasing effects that arise from sampling non-bandlimited analytical objects below the Nyquist rate.

We computed the confocal point spread function from a

theoretical model of the confocal image formation, based on electromagnetic diffraction theory (van der Voort & Brakenhoff, 1990). This model takes important microscopical parameters, such as the finite-size pinhole, high apertures and polarization effects into account; lens aberrations are not modelled.

The performance of the restoration algorithms is measured as a function of the signal-to-noise ratio (*SNR*), which we define as

$$SNR = \frac{E}{\epsilon}$$

(23)

with ϵ the total power of the noise and E the total power of the object. The simulated images are distorted by Poisson noise. The noise is generated by using the intensity of the convolved spheres as averages of a spatially variant Poisson process (Press *et al.*, 1992). We have varied the signal-to-noise ratio of the simulated images by changing the photon-conversion efficiency. For a Poisson process, the variance σ^2 equals the mean, so the noise power in the image is equal to

$$\epsilon = \sigma^2 = c^{-1}(\tfrac{4}{3}\pi r^3 I_0 + VI_b)$$

(24)

with c the photon-conversion efficiency (photons/ADU), V the image volume (μm^3), I_0 the average sphere intensity (ADU) and I_b the average intensity of the background (ADU), and r radius of the sphere (μm). Using (23) and (24) the photon-conversion can now be found with

$$c = \frac{SNR \cdot (\tfrac{4}{3}\pi r^3 I_0 + VI_b)}{\tfrac{4}{3}\pi r^3 I_0^2}.$$

(25)

For our simulations, we have selected microscope parameters corresponding to typical working conditions: a numerical aperture of 1·3, a refractive index of the lens immersion oil of 1·515, an excitation wavelength of 479 nm, an excitation/emission ratio of 0·9 and a pinhole size of 282 nm. These conditions result in a lateral sampling distance of 46·0 nm and an axial sampling distance of 162·4 nm, when the images are sampled at the Nyquist frequency.

3.2. Iterative optimization: where to start and when to stop

All three investigated restoration algorithms need a first guess to start their iterations. We have used the measured image m as a first estimate $\hat{f}_0$ to start the Richardson–Lucy and ICTM algorithms. To start the Carrington algorithm we have set c_0 to zero, as proposed by Carrington (1990).

In principle, one can continue to generate new estimates of $\hat{f}$ until the optimum of functional is found by the restoration algorithm. In practice, this procedure is undesirable. Experiments (Holmes & Liu, 1991) show that the likelihood of a Richardson–Lucy estimate increases logarithmically as a function of the number of iterations. This growth makes the search for the maximum of the

likelihood function extremely computationally expensive. We have therefore used a threshold (we have chosen 0·001%) on the change of the functional ($(f_{k+1} - f_k)/f_k$) to stop the iteration. This criterion can be seen as a speed-of-convergence criterion, since it is a threshold on the slope of the change of the functional as function of the number of iterations. The functionals of the methods considered have been shown to converge as an exponential function of the number of iterations to their optimum (Holmes & Lui, 1991; Lagendijk & Biemond, 1991). Therefore, the threshold determines how near to the optimum the algorithms are stopped.

3.3. Estimation of the regularization parameter

The Tikhonov functional (10) shows that the restoration results will be smoother for higher values of the regularization parameter λ (more regularization). Lower values of λ will result in 'crisper' results that are, in general, more sensitive to the noise in the image. Previous work (van Kempen et al., 1996) showed that when the regularization parameter λ equals $1/SNR$, the result of the ICTM algorithm is too smooth. It is therefore of great importance to have a reliable method for determining the λ parameter used in the ICTM and Carrington algorithms. In Galatsanos & Katsaggelos (1992) the methods of constrained least-squares, generalized cross-validation and maximum likelihood are described to determine λ for the closed formed Tikhonov–Miller solution (11). In this paper, we compare these three methods of determining λ with λ equal to $1/SNR$ for the ICTM and Carrington algorithms.

The method of constrained least-squares (CLS) finds a λ_{CLS} such that the following equation holds:

$$\|M(\omega) - H(\omega)\hat{F}(\omega)\|^2 = \|M(\omega) - H(\omega)A(\omega)G(\omega)\|^2 = \epsilon \tag{26}$$

with ϵ the total noise power and $A(\omega)$ given in (11). As the total noise power ϵ is determined by the variance of the noise, this method requires knowledge of the noise variance. The method of generalized cross-validation (GCV) can be derived from the leave one out principle, and can be expressed in the discrete Fourier domain as (Galatsanos & Katsaggelos, 1992)

$$GCV(\lambda) = \frac{\displaystyle\sum_\Omega \frac{\lambda^2 |R(\omega)|^4 |M(\omega)|^2}{(|H(\omega)|^2 + \lambda|R(\omega)|^2)^2}}{\left(\displaystyle\sum_\Omega \frac{\lambda|R(\omega)|^2}{|H(\omega)|^2 + \lambda|R(\omega)|^2}\right)^2} \tag{27}$$

with λ_{GCV} the value of λ that minimizes (27). This minimum can be found without prior knowledge of the noise variance (Reeves & Mersereau, 1992).

An alternative method to determine λ without prior knowledge of the noise variance has been named the maximum likelihood method by Galatsanos & Katsaggelos (1992). It is based on a stochastic approach, which assumes $\int_X r(x - \chi)f(\chi)\,d\chi$ and the noise to be Gaussian distributed (this can be achieved with a proper choice of the regularization filter $r(x)$, see Galatsanos & Katsaggelos (1992)). The derived maximum likelihood function can be evaluated in the discrete Fourier domain with (Galatsanos & Katsaggelos (1992))

$$ML(\lambda) = \frac{\displaystyle\sum_\Omega \frac{\lambda|R(\omega)|^2 |M(\omega)|^2}{|H(\omega)|^2 + \lambda|R(\omega)|^2}}{\displaystyle\prod_\Omega \left(\frac{\lambda|R(\omega)|^2}{|H(\omega)|^2 + \lambda|R(\omega)|^2}\right)^{1/\Omega}} \tag{28}$$

with Ω number of Fourier coefficients. In our experiments, we have determined the value of the regularization parameter, using one of these three methods, before starting the ICTM and the Carrington algorithm.

3.4. Restoration of spheres

This experiment compares the capability of the Richardson–Lucy, the ICTM and the Carrington algorithm to restore spheres convolved with a confocal PSF and distorted by Poisson noise. The performance of the algorithms is measured with the MSE and I-divergence measure as a function of the SNR. We have generated spheres with a radius of $1·0\,\mu m$, an object intensity of 200·0 and a background of 40·0. The images are $128 \times 128 \times 32$ pixels in size, the SNR ranges from 1·0 to 256·0 (0·0–24·2 dB). Using (25) this corresponds to 9·5–2435 photons per voxel in the object and 1·9–487 photons per voxel in the background.

In Fig. 1, the value of the regularization parameter is plotted as a function of the SNR as determined by $1/SNR$, the CLS, GCV or ML method.
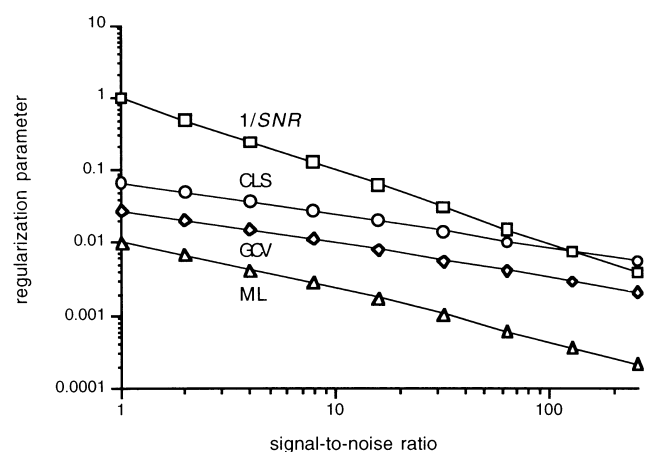

Fig. 1. The value of the regularization parameter as a function of the signal-to-noise ratio. The regularization parameter was determined by the CLS, GCV or ML method, or set at $1/SNR$.
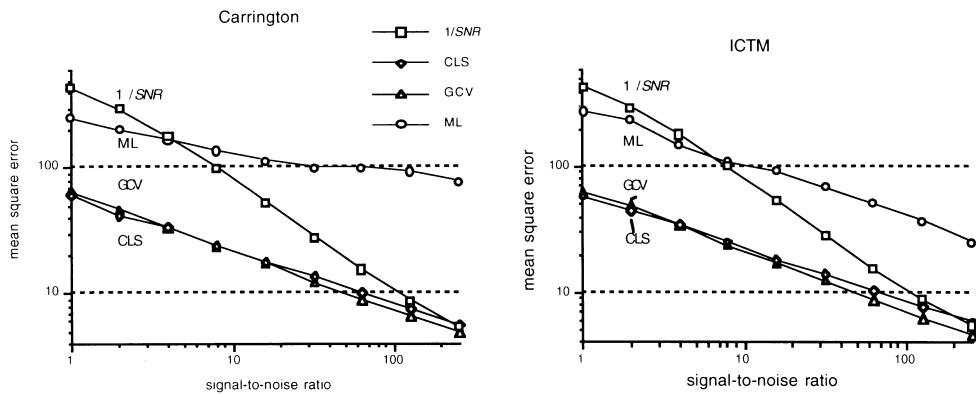
Fig. 2. The value of the mean square error of the restoration result of the Carrington and ICTM algorithms. The regularization parameter was determined with the CLS, GCV and ML method and by setting it to 1/*SNR*.
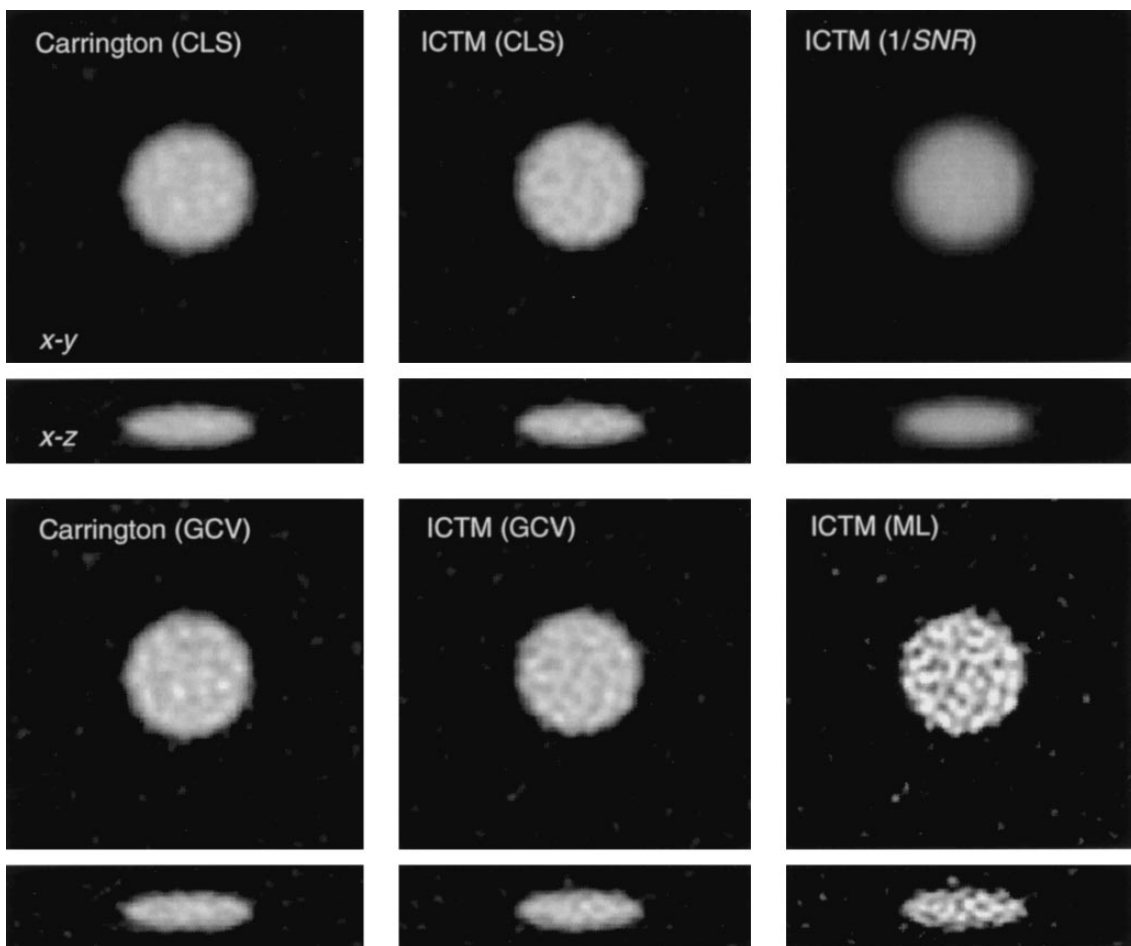

Fig. 3. The restoration results of the Carrington and ICTM algorithms with the regularization parameter determined by 1/*SNR* and with the CLS, GCV and ML methods. The pictures show the centre $x-y$ and $x-z$ slices of an image ($128 \times 128 \times 32$ pixels in size) of a sphere with radius $1 \cdot 0 \, \mu$m and a SNR of $16 \cdot 0$. The object and its confocal image are shown in Fig. 5. The images are displayed with eight-bit grey-scale resolution, without stretching the intensities.
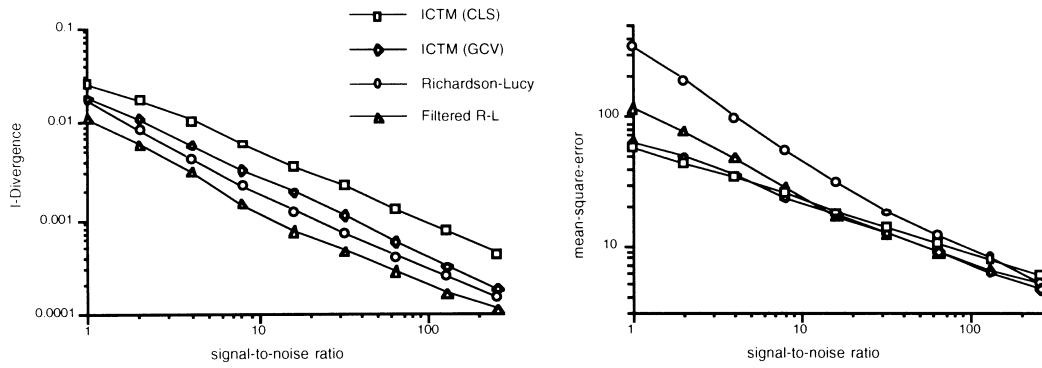
Fig. 4. The I-divergence (left) and mean-square-error (right) performance of the ICTM algorithm and the Richardson–Lucy algorithm. ICTM's regularization parameter has been determined by the CLS and the GCV algorithm. The Richardson–Lucy algorithm is tested with and without noise suppression by Gaussian prefiltering
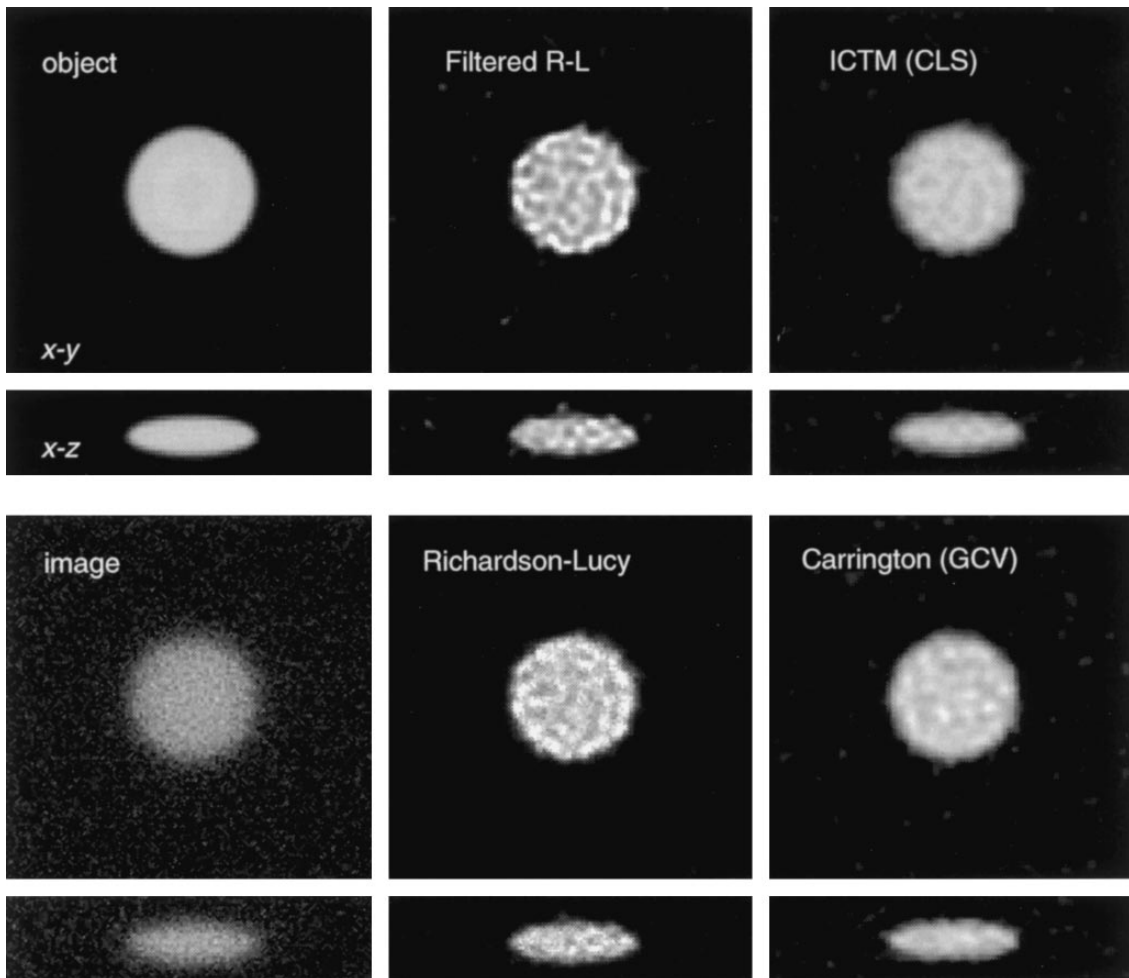

Fig. 5. Restoration of spheres. From left to right we show the centre $x-y$ and $x-z$ slices of the object and its confocal image (left), the results of Richardson–Lucy and filtered R-L (middle), and the results of the ICTM (CLS) and Carrington (GCV) algorithms. The images are $128 \times 128 \times 32$ pixels in size, the radius of the sphere is $1 \cdot 0 \, \mu m$ and SNR is $16 \cdot 0$. The images are displayed with eight-bit grey-scale resolution, without stretching the intensities. The background has been subtracted from the measured image, to ease the comparison with the other images.
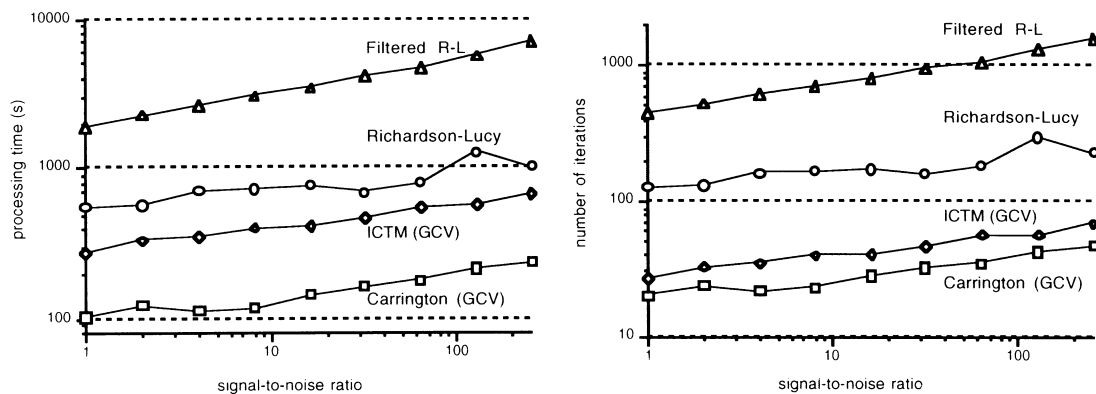
Fig. 6. The processing time (left) and the number of iterations (right) needed by the Richardson–Lucy, filtered R-L, Carrington and ICTM algorithms as a function of the SNR. The regularization parameter of the Carrington and ICTM algorithm is determined by the GCV method.

Figure 2 shows that the results of the ICTM and Carrington algorithms converge to the same solution in the mean square error sense. Furthermore, it shows the large influence of the regularization parameter on the results. Large ($1/SNR$) as well as small values of λ (ML method) will result in poor restoration results. The restoration results of the Carrington and ICTM algorithms are shown in Fig. 3. The regularization parameter is determined by $1/SNR$, the CLS, GCV or ML methods. Figure 3 shows again the similarity of the Carrington and ICTM results with λ determined by the CLS or GCV method. Furthermore, it shows that the ICTM result with λ set to $1/SNR$ is too smooth, whereas the result with the ML method shows noise adaptation caused by the small value of the λ, reducing the regularization.

Figure 4 shows the I-divergence and mean square error performance of the ICTM algorithm and the Richardson–Lucy algorithm. We have employed the CLS and GCV methods for determining the regularization parameter of the ICTM algorithm. We show the results of the Richardson–Lucy algorithm with and without noise suppression by Gaussian prefiltering (as described in Section 2). We have used a Gaussian with a sigma of 1 pixel in the lateral direction and of 2 pixels in the axial direction. The Richardson–Lucy algorithm outperforms the ICTM algorithm with respect to the I-divergence criterion, but the ICTM algorithm performs better than Richardson–Lucy in mean-square-error sense for low *SNR*. Figure 5 shows centre $x–y$ and $x–z$ slices of the sphere, the confocal image as well as the results of the four restoration algorithms.

Figure 6 shows the processing time and the number of iterations needed by the Richardson–Lucy, filtered R-L, Carrington and ICTM algorithms as a function of the SNR. The λ of the Carrington and ICTM algorithms is determined by the GCV method. The processing time was measured on a Sun UltraServer 1 computer, equipped with an UltraSparc 1 CPU running at 143 MHz and 160 MB main memory (Sun Microsystems, Mountain View, CA).

This difference in execution time is explained by two factors: the number of iterations and the number of Fourier transforms per iteration. The Richardson–Lucy algorithm needs about six times more iterations than the Carrington algorithm before our stop criterion is met. The filtered Richardson–Lucy algorithm needs about 25 times more iterations; the ICTM algorithm needs on average 1·5 times more iterations. Secondly, both the Richardson–Lucy and the Carrington algorithm need four Fourier transforms per iteration. The ICTM algorithm, on the other hand, needs two Fourier transforms per iteration, and two transforms to start an iterative search for β_k, which needs another transform per subiteration. Our implementation of the ICTM needs about 4–6 iterations to determine β_k. Therefore, the ICTM algorithm needs, in total, 2–2·5 times more Fourier transforms per iteration than the Carrington algorithm. It should be noted that the number of iterations of the Richardson–Lucy type algorithms could be reduced considerably if an acceleration method as proposed by Holmes & Liu (1991) is adapted to work properly in the presence of background term $b(x)$.

3.5. Improvement of quantitative analysis

This experiment tests the ability of the Richardson–Lucy, ICTM and Carrington algorithms to improve the quantitative measurement of the total amount of fluorescence inside one sphere in the vicinity of a second sphere as a function of the distance between the spheres. We have generated one sphere with a radius of 0·4 μm and an intensity of 200·0 and a second sphere with a radius of 0·6 μm and an intensity of 50·0 (the background remained 40·0). A fixed distance of 0·05 μm is used in the lateral direction and is varied in the axial direction from 0·0 to 2·0 μm. We chose to vary the axial distance as most of the blurring occurs in this direction. The total amount of fluorescence inside the spheres is measured before convolving it with the confocal PSF, to obtain the ground truth, and before and after
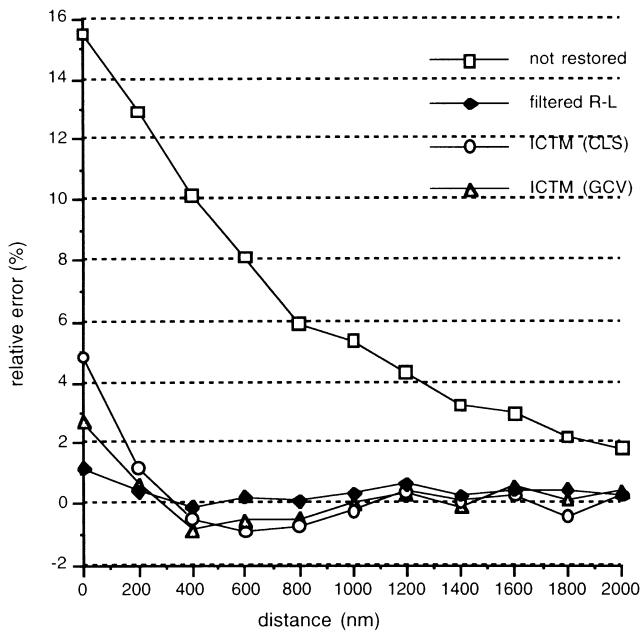
Fig. 7. The relative error of the total intensity of one sphere as a function of axial distance to a second sphere. The intensity is measured before and after restoration with the filtered Richardson–Lucy or the ICTM algorithm (using GCV or CLS to determine λ).

restoration. As described in Section 2.3 we measured the total amount of fluorescence inside a sphere by assigning the intensity of a pixel to the sphere with the shortest distance from the pixel to the sphere's surface. In this experiment we used a sample distance of 46·0 nm in both the lateral and the axial direction. The images have a size of $64 \times 64 \times 128$ pixels, and a SNR of 16·0 was used. Figure 7 shows the error of the total intensity measured before and after restoration relative to the total intensity measured

before convolving the spheres with the confocal PSF. Two spheres (with an axial distance of 0·2 μm) and their confocal image are shown in Fig. 8 together with the restoration results of the filtered Richardson–Lucy and the ICTM algorithms (using the GCV method for determining λ).

4. Discussion and conclusions

We have compared the performance of the Richardson–Lucy, ICTM and Carrington restoration algorithms applied to 3-D confocal test images. All algorithms greatly reduce diffraction-induced distortions of these images. The reader should be aware that we have generated 'simple' objects: spheres with a constant intensity, which *could* make these images more beneficial for imposed regularization than images of more 'complex' objects. For example, texture or cell cores, generally found in biological images, have not been modelled. Furthermore, we have used a theoretical model of the confocal PSF, which does not model lens aberrations and refractive-index-induced diffraction, which could influence the performance of the restoration methods in practical situations.

From our experiments we have found that the differences between the Carrington and ICTM algorithms are small in the mean-square-error sense compared with the Richardson–Lucy algorithm, despite the different incorporation of the non-negativity constraint in the two algorithms which results in two different strategies to find the minimum of the Tikhonov functional. We found that a small value of the regularization parameter (determined by the ML method) results in poor and different results for these two algorithms. A proper value of the regularization parameter was found by the constrained least-squares (CLS) method and the method of generalized cross-validation (GCV). Setting the
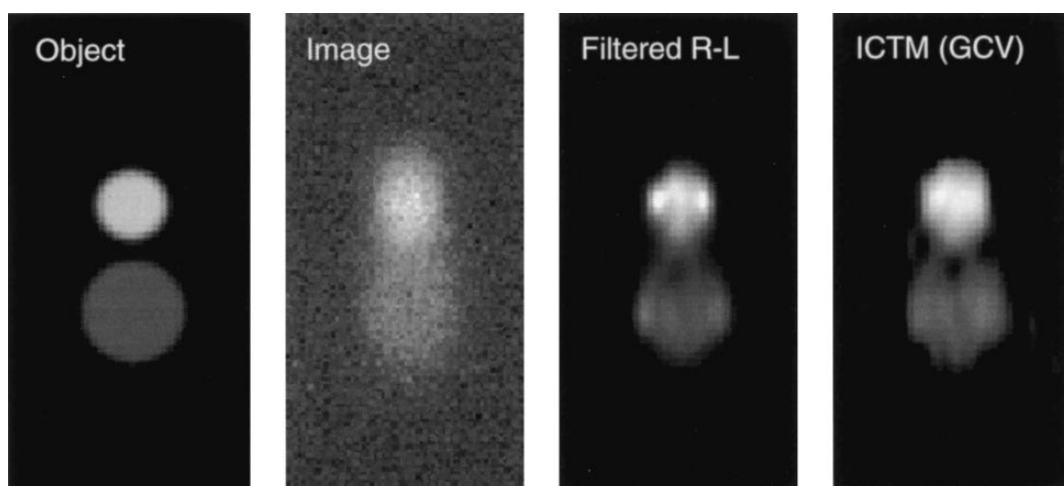

Fig. 8. Two spheres. From left to right we show the centre x–z slice of the object, image and of the restoration results of the filtered R-L and ICTM algorithms. The axial distance between the two spheres is 0·2 μm.

regularization parameter equal to the theoretical value of $1/SNR$, as we have done in a previous experiment (van Kempen *et al.*, 1996), results in a solution which is too smooth. The CLS method for determining λ needs prior knowledge of the noise variance. In our simulations this was easily derived from the SNR we specified. In practice, however, this is not possible. The average noise variance can then be estimated from the high-frequency band of the image spectrum which is dominated by noise.

We have found that the results of the Richardson–Lucy algorithm improve in I-divergence as well as mean-square-error sense by prefiltering both the image and the PSF with a small Gaussian. This Gaussian prefiltering will reduce noise sensitivity by suppressing those parts of the image spectrum that do not contain any signal information or which are dominated by noise. A disadvantage of this filtering is the large increase in the number of iterations needed. However, we are currently investigating the application of the acceleration method of Holmes & Liu (1991) for the Richardson–Lucy algorithm that incorporates the background term $b(x)$.

The ICTM and Carrington algorithms perform better in mean-square-error sense than the (improved) Richardson–Lucy for images with a low signal-to-noise ratio (given a proper choice of the regularization parameter). This can be caused by the fact that the ICTM and Carrington algorithms are designed to minimize the mean-square-error or by the regularization they incorporate. However, it must be stressed that, under low SNR conditions, the MSE criterion is not an appropriate difference measure for images used which were hampered by Poisson noise.

Although the investigated image restoration algorithms minimize the *MSE* or I-divergence measure, these measures are not conclusive when the goal of image restoration is to improve the quantitative analysis of microscopic images. We have shown that the measurement of the total amount of fluorescence inside a sphere in the vicinity of another sphere is greatly improved by applying image restoration prior to the analysis. In this particular experiment, the filtered Richardson–Lucy algorithm produces better results than the ICTM algorithm.

Acknowledgments

References

Art, J. (1995) Photon detectors for confocal microscopy. *Handbook of Biological Confocal Microscopy* (ed. by J. B. Pawley), pp. 183–196. Plenum Press, New York.

Carrington, W.A. (1990) Image restoration in 3D microscopy with limited data. *Bioimaging and Two-Dimensional Spectroscopy Proc. SPIE* (ed. by L. C. Smith), Vol. 1205, pp. 72–83. SPIE.

Carrington, W.A., Lynch, R.M., Moore, E.M., Isenberg, G., Fogarty, K.E. & Fay F.S. (1995) Superresolution three-dimensional images of fluorescence in cells with minimal light exposure. *Science*, **268**, 1483–1487.

Csiszár, I. (1991) Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Ann. Statist.* **19**, 2032–2066.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc. B*, **39**, 1–37.

Galatsanos, N.P. & Katsaggelos, A.K. (1992) Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Trans. Image Processing*, **1**, 322–336.

Holmes, T.J., Bhattacharyya, S., Cooper, J.A., Hanzel, D. & Krishnamurti, V. (1995) Light microscopic images reconstructed by maximum likelihood deconvolution. *Handbook of Biological Confocal Microscopy* (ed. by J. B. Pawley), pp. 389–402. Plenum Press, New York.

Holmes, T.J. & Liu, Y.H. (1991) Acceleration of maximum-likelihood image restoration for fluorescence microscopy and other noncoherent imagery. *J. Opt. Soc. Am. A*, **8**, 893–907.

Joshi, S. & Miller, M.I. (1993) Maximum *a posteriori* estimation with good roughness for 3-D optical-sectioning microscopy. *Opt. Soc. Am. A*, **10**, 1078–1085.

van Kempen, G.M.P., van der Voort, H.T.M., Bauman, J.G.J. & Strasters, K.C. (1996) Comparing maximum likelihood estimation and constrained Tikhonov–Miller restoration. *IEEE Eng. Med. Biol.* **15**, 76–83.

Lagendijk, R.L. & Biemond, J. (1991) *Iterative Identification and Restoration of Images.* Kluwer, Boston.

Pawley, J.B. (1990) Fundamental limits in Confocal microscopy. *Handbook of Biological Confocal Microscopy* (ed. by J. B. Pawley), pp. 15–26. Plenum Press, New York.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. (1992) *Numerical Recipes in C*, 2nd edn. Cambridge University Press, Cambridge.

Reeves, S.J. & Mersereau, R.M. (1992) Blur identification by the method of generalized cross-validation. *IEEE Trans. Image Processing*, **1**, 301–311.

Wijnaends van Resandt, R., Marsman, H.B.J., Kaplan, R., Davoust, J., Stelzer, E.K.H. & Stricker, R. (1985) Optical fluorescence microscopy in 3 dimensions: microtomoscopy. *J. Microsc.* **138**, 29–34.

Richardson, W.H. (1972) Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* **62**, 55–59.

Shaw, P.J. & Rawlins, D.J. (1991) The point-spread function of a

confocal microscope: its measurement and use in deconvolution of 3-D data. *J. Microsc.* **163,** 151–166.

Shepp, L.A. & Vardi, Y. (1982) Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging,* **MI-1,** 113–121.

Snyder, D.L., Hammoud, A.M. & White, R.L. (1993) Image recovery from data acquired with a charge-coupled-device camera. *J. Opt. Soc. Am. A,* **10,** 1014–1023.

Snyder, D.L. & Miller, M.I. (1991) *Random Point Processes in Time and Space.* Springer Verlag, Berlin.

Snyder, D.L., Schutz, T.J. & O'Sullivan, J.A. (1992) Deblurring subject to nonnegative constraints. *IEEE Trans. Signal Processing.* **40,** 1143–1150.

Tikhonov, A.N. & Arsenin, V.Y. (1977). *Solutions of Ill-Posed Problems.* Wiley, New York.

Vardi, Y., Shepp, L.A. & Kaufman, L. (1985) A statistical model for positron emission tomography. *J. Am. Statist. Assoc.* **80,** 8–35.

van Vliet, L.J. (1993) *Grey-scale Measurements in Multi-Dimensional Digitized Images.* Delft University Press, Delft.

van der Voort, H.T.M. & Brakenhoff, G.J. (1990) 3-D image formation in high-aperture fluorescence confocal microscopy: a numerical analysis. *J. Microsc.* **158,** 43–54.

van der Voort, H.T.M. & Strasters, K.C. (1995) Restoration of confocal images for quantitative image analysis. *J. Microsc.* **178,** 165–181.