Journal of Microscopy, Vol. 198, Pt 1, April 2000, pp. 63–75.
Received 31 March 1999; accepted 28 October 1999

The influence of the regularization parameter and the first estimate on the performance of Tikhonov regularized non-linear image restoration algorithms

G. M. P. VAN KEMPEN & L. J. VAN VLIET
Pattern Recognition Group, Faculty of Applied Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

Key words. First estimate, image restoration, iterative restoration algorithms, regularization parameter.

Summary

This paper reports studies on the influence of the regularization parameter and the first estimate on the performance of iterative image restoration algorithms. We discuss regularization parameter estimation methods that have been developed for the linear Tikhonov–Miller filter to restore images distorted by additive Gaussian noise. We have performed experiments on synthetic data to show that these methods can be used to determine the regularization parameter of non-linear iterative image restoration algorithms, which we use to restore images contaminated by Poisson noise. We conclude that the generalized cross-validation method is an efficient method to determine a value of the regularization parameter close to the optimal value. We have also derived a method to estimate the regularization parameter of a Tikhonov regularized version of the Richardson–Lucy algorithm.

These iterative image restoration algorithms need a first estimate to start their iteration. An obvious and frequently used choice for the first estimate is the acquired image. However, the restoration algorithm could be sensitive to the noise present in this image, which may hamper the convergence of the algorithm. We have therefore compared various choices of first estimates and tested the convergence of various iterative restoration algorithms. We found that most algorithms converged for most choices, but that smoothed first estimates resulted in a faster convergence.

Correspondence: G. M. P. van Kempen, Central Analytical Sciences, Spectroscopy & Microscopy Unit, Unilever Research Vlaardingen, Olivier van Noortlaan 120, 3133 AT Vlaardingen, The Netherlands.

[1] Although mathematically equivalent to the maximum a priori (MAP) approach used here, the maximum penalized likelihood method (MPL) employs the regularization term as a means to ameliorate some undesired behaviour of the solution instead of an, often hard to obtain, a priori probability distribution of the scene under observation.

1. Introduction

The goal of image restoration is to 'undo' the blurring and noise artefacts that are imposed on the image during image acquisition. In other words, image restoration algorithms try to restore the 'original' image from the acquired image by deconvolving the blurring imposed by the point spread function (PSF) and by reducing the noise imposed by the image recording.

The performance of non-linear iterative image restoration algorithms is highly dependent on input parameters of these algorithms, such as the PSF, background value, regularization parameter, first estimate and stop criterion. For example, the dependence of the performance of these algorithms on the estimation of the PSF and the background estimation have been shown before (van der Voort & Strasters, 1995; van Kempen, 1999; van Kempen *et al.*, 1997). We studied the performance of non-linear iterative regularized image restoration algorithms as a function of the regularization parameter and the first estimate. The first estimate is a first estimate of the 'original image' that iterative restoration algorithms require as input.

We studied regularized image restoration algorithms like the iterative constrained Tikhonov–Miller algorithm (ICTM) (Lagendijk & Biemond, 1991; van der Voort & Strasters, 1995), the Carrington algorithm (Carrington, 1990; Carrington *et al.*, 1995) and the Tikhonov regularized Richardson–Lucy algorithm, which we refer to as the RL–Conchello algorithm (Conchello & McNally, 1996). These algorithms balance the fit of their restoration result to the acquired image with an a priori model of the restoration result[1] (Andrews & Hunt, 1977). This balance is determined by the regularization parameter, which determines to what extent the restoration result is governed by the noisy recorded data or by the prior. A large value for the regularization parameter results in a stronger influence of

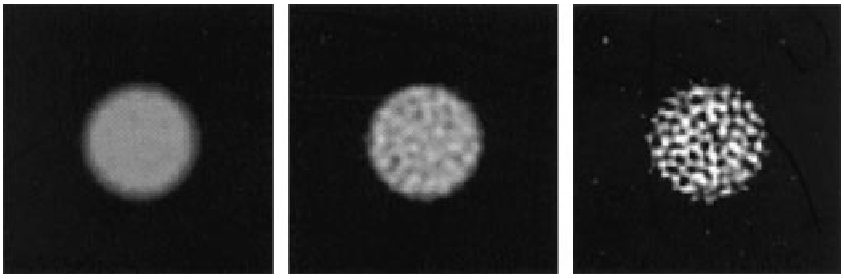


Fig. 1. Restoration of a sphere using the ICTM algorithm for a high value (0.1) (left), a low value (0.00001) (right) and a well-chosen value (0.006) for the regularization parameter.

the regularization prior on the restoration result, yielding in general a smooth result, whereas a low value of the regularization parameter will make the restoration algorithms more sensitive to the noise in the acquired image (see Fig. 1). The regularization parameter therefore is an important factor in the performance of these image restoration algorithms. We investigate the influence of the regularization parameter on the performance of various restoration algorithms and compare various methods for determining the regularization parameter for the ICTM and Carrington algorithm. Furthermore, we show how the regularization parameter for the RL–Conchello algorithm can be estimated.

This paper presents a more thorough investigation of the influence of the regularization parameter of the performance of non-linear restoration algorithms than previous work (van Kempen *et al.*, 1997). It compares a computed regularization parameter with the value that is optimal in mean square error sense (a value that can only be obtained in a simulation experiment). Furthermore, this paper presents a novel method for the estimation of the regularization parameter for the RL–Conchello algorithm.

The algorithms mentioned are iterative, non-linear image restoration algorithms. These iterative algorithms need a first estimate to start their iterations. A commonly used choice is the acquired image itself. Although this image is often close enough to the end result to guarantee a convergence of the iteration, the acquired image contains noise, to which the algorithms can be sensitive. Therefore, it is useful to compare this choice of first estimate to alternatives. We have investigated smoothed versions of the acquired image as first estimate as well as the result of linear Tikhonov restoration. This latter choice will in general be much closer to the final restoration result, especially for confocal imaging, therefore promising (much) faster convergence.

In the next section we formulate a mathematical description of the image formation in a confocal fluorescence microscope and use it to discuss linear and non-linear image restoration techniques. In section 3, we review various methods to estimate the regularization parameter of the Tikhonov functional. The next section compares the performances of the ICTM and Carrington algorithms as function of these regularization parameter estimations. In section 5, we derive a method to estimate the regularization parameter for the RL–Conchello algorithm. Section 6 discusses various first estimates and shows how the convergence of various iterative restoration algorithms depends on the choice of first estimate.

2. Image restoration

2.1. Classical image restoration: the Tikhonov–Miller restoration filter

We assume that the image formation in a confocal fluorescence microscope can be modelled as a linear translation-invariant system distorted by noise

$$m(x,y,z) = N\,(h(x,y,z) \otimes f(x,y,z) + b\,(x,y,z)) \tag{1}$$

In this equation f represents the input signal, h the point spread function, b a background signal (which we assume to be constant), N a general noise distortion function, and m the recorded fluorescence image. For scientific-grade light detectors N is dominated by Poisson noise.

In classical image restoration, the signal-dependent Poisson noise is approximated by additive Gaussian noise. Using this additive Gaussian noise model for N we rewrite (1) as

$$g(x,y,z) = m\,(x,y,z) - b\,(x,y,z) = h(x,y,z) \otimes f(x,y,z) + n(x,y,z) \tag{2}$$

After sampling Eq. (2) becomes

$$g[x,y,z] = \sum_{i=1}^{M_x}\sum_{j=1}^{M_y}\sum_{k=1}^{M_z} h[x-i, y-j, z-k]\,f\,[i,j,k] + n[x,y,z] \tag{3}$$

with M_x, M_y and M_z the number of sampling points in the x, y and z dimension, respectively. For convenience we will adopt a matrix notation

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \tag{4}$$

where the vectors **f**, **g** and **n** of length M ($M = M_x + M_y + M_z$) denote the object, its image and the additive Gaussian noise, respectively. The MxM matrix **H** is the blurring matrix representing the PSF of the microscope.

The Tikhonov–Miller filter, a classical image restoration filter, is a convolution filter operating on the measured image. It can be written as

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{g} \tag{5}$$

with $\mathbf{W}$ the linear restoration filter and $\hat{\mathbf{f}}$ its result. The Tikhonov–Miller filter is derived from a least squares approach, which is based on minimizing the squared difference between the acquired image and a blurred estimate of the original object,

$$\|\mathbf{H}\hat{\mathbf{f}} - \mathbf{g}\|^2 \tag{6}$$

However, a direct minimization of Eq. (6) will produce undesired results, as it does not take into account the (high) frequency components of $\hat{\mathbf{f}}$ that are set to zero by the convolution with $\mathbf{H}$.

Finding an estimate $\hat{\mathbf{f}}$ from Eq. (6) is known as an ill-posed problem (Tikhonov & Arsenin, 1977). To address this issue Tikhonov defined the regularized solution $\hat{\mathbf{f}}$ of Eq. (4) the one that minimizes the well-known Tikhonov functional (Tikhonov & Arsenin, 1977)

$$\Phi(\hat{\mathbf{f}}) = \|\mathbf{H}\hat{\mathbf{f}} - \mathbf{g}\|^2 + \lambda\|\mathbf{R}\hat{\mathbf{f}}\|^2 \tag{7}$$

with $\|\cdot\|^2$ the Euclidean norm. In image restoration λ is known as the regularization parameter and $\mathbf{R}$ as the regularization matrix. The Tikhonov functional consists of a mean square error fitting criterion and a stabilizing energy bound which penalizes solutions of $\hat{\mathbf{f}}$ that oscillate wildly due to spectral components which are dominated by noise. The minimum of Φ can be found by solving

$$\nabla_{\hat{\mathbf{f}}}\Phi(\hat{\mathbf{f}}) = 2\mathbf{H}^T(\mathbf{H}\hat{\mathbf{f}} - \mathbf{g}) + 2\lambda\mathbf{R}^T\mathbf{R}\hat{\mathbf{f}} = 0 \tag{8}$$

which yields the well-known Tikhonov–Miller (TM) solution $\mathbf{W}_{TM}$ (Miller, 1970; Tikhonov & Arsenin, 1977)

$$\mathbf{W}_{TM} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{R}^T\mathbf{R})^{-1}\mathbf{H}^T \tag{9}$$

The linear nature of the Tikhonov–Miller makes it incapable of restoring frequencies for which the PSF has a zero response. Furthermore, linear methods cannot restrict the domain in which the solution should be found. This property is a major drawback as the intensity of an imaged object represents light energy, which is non-negative.

The ICTM algorithm, the Carrington algorithm and the Richardson–Lucy algorithm are frequently used in fluorescence microscopy (Holmes, 1988; Carrington, 1990; van der Voort & Strasters, 1995; Conchello & McNally, 1996; van Kempen *et al.*, 1997; van Kempen, 1999; Verveer *et al.*, 1999). These iterative, non-linear algorithms tackle the above mentioned problems in exchange for a considerable increase in the computational complexity.

2.2. Constrained Tikhonov restoration

2.2.1. The iterative constrained Tikhonov–Miller algorithm. The ICTM algorithm (Lagendijk & Biemond, 1991; van der Voort & Strasters, 1995) finds the minimum of Eq. (7) using the method of conjugate gradients (Press *et al.*, 1992). Setting the negative intensities to zero after each iteration implements the non-negativity constraint.

2.2.2. The Carrington algorithm. Like the ICTM algorithm, the Carrington algorithm (Carrington, 1990; Carrington *et al.*, 1995) minimizes the Tikhonov functional under the constraint of non-negativity. However, the algorithm is based on a more solid mathematical foundation. The derivation of the Carrington algorithm as given in this section is based on Verveer (personal communication, 1995).

Given the Tikhonov functional Eq. (7) (without the regularization matrix) and its derivative

$$\frac{1}{2}\nabla_{\hat{\mathbf{f}}}\Phi(\hat{\mathbf{f}}) = \mathbf{H}^T(\mathbf{H}\hat{\mathbf{f}} - \mathbf{g}) + \lambda\hat{\mathbf{f}} \tag{10}$$

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

$$\nabla_{\hat{\mathbf{f}}}\Phi_i = 0 \text{ and } \hat{\mathbf{f}}_i > 0 \quad \text{or} \quad \nabla_{\hat{\mathbf{f}}}\Phi_i \geq 0 \text{ and } \hat{\mathbf{f}}_i = 0 \tag{11}$$

Then from Eq. (10) we find

$$\nabla_{\hat{\mathbf{f}}}\Phi = 0 \rightarrow \hat{\mathbf{f}} = \frac{1}{\lambda}\mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}) \equiv \mathbf{H}^T\mathbf{c} \tag{12}$$

From these conditions we see that:

$$\hat{\mathbf{f}} = \mathbf{H}^T\mathbf{c} \textit{ where } \mathbf{H}^T\mathbf{c} > 0 \quad \textit{and} \quad \hat{\mathbf{f}} = 0 \textit{ where } \mathbf{H}^T\mathbf{c} \leq 0 \tag{13}$$

This can be written as

$$\hat{\mathbf{f}} = \max(0, \mathbf{H}^T\mathbf{c}) = P(\mathbf{H}^T\mathbf{c}) \tag{14}$$

On the set where $\mathbf{H}^T\mathbf{c} > 0$ we then obtain after inserting (14) into (8) that

$$\nabla_{\hat{\mathbf{f}}}\Phi = \mathbf{H}^T(\mathbf{H}P(\mathbf{H}^T\mathbf{c}) - \mathbf{g} + \lambda\mathbf{c}) = 0 \tag{15}$$

This is equivalent to minimizing the functional Ψ (on the set $\mathbf{H}^T\mathbf{c} > 0$)

$$\Psi(\mathbf{c}) = \frac{1}{2}\|P(\mathbf{H}^T\mathbf{c}\|^2 - \mathbf{c}^T\mathbf{g} + \frac{1}{2}\lambda\|\mathbf{c}\|^2 \tag{16}$$

Since Ψ is strictly convex and twice continuously differentiable, a conjugate gradient algorithm can be used to minimize Ψ.

2.3. Maximum likelihood restoration: the Richardson–Lucy algorithm

By contrast with the two algorithms discussed previously the Richardson–Lucy is not derived from the image formation model (4) which assumes additive Gaussian noise. Instead the general noise distortion function N is assumed to be dominated by Poisson noise.

A fluorescence object can be modelled as a spatially inhomogeneous Poisson process $\mathbf{F}$ with an intensity function $\mathbf{f}$ (Snyder & Miller, 1991),

$$P(\mathbf{F}_i|\mathbf{f}_i) = \frac{\mathbf{f}_i^{\mathbf{F}_{i_e}-\mathbf{f}_i}}{\mathbf{F}_i!} \tag{17}$$

The image formation of such an object by a fluorescence microscope can be modelled as a translated Poisson process (Snyder & Miller, 1991). This process models the transformation of $\mathbf{F}$ into a Poisson process $\mathbf{m}$ subjected to a conditional probability density function $\mathbf{H}$,

$$E[\mathbf{m}] = \mathbf{Hf} + \mathbf{b} \tag{18}$$

with $\mathbf{b}$ the mean of an independent (background) Poisson process. The log likelihood function of such a Poisson process is given by (Snyder & Miller, 1991)

$$L(\mathbf{f}) = -\sum \mathbf{Hf} + \mathbf{m}^T \ln(\mathbf{Hf} + \mathbf{b}) \tag{19}$$

where we have dropped all terms that are not dependent on $\mathbf{f}$. The maximum of the likelihood function L can be found iteratively using the EM algorithm as described by Dempster *et al.* (1977). This iterative algorithm was first used by Vardi *et al.* (1985) in emission tomography. Holmes (1988) introduced the algorithm to microscopy. Applying the EM algorithm to Eq. (19) yields (Shepp & Vardi, 1982; Holmes, 1988; Snyder *et al.*, 1993; van Kempen *et al.*, 1997; van Kempen, 1999)

$$\hat{\mathbf{f}}^{k+1} = \hat{\mathbf{f}}^k (\mathbf{H}\hat{\mathbf{f}}^k + \mathbf{b})^{-1} \mathbf{H}^T \mathbf{m} \tag{20}$$

The EM algorithm ensures a non-negative solution when a non-negative initial guess $\hat{\mathbf{f}}^0$ 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 for finding the maximum likelihood estimator of a translated Poisson process (often referred to as EM-MLE) is identical to the Richardson–Lucy algorithm (Richardson, 1972).

The Richardson–Lucy algorithm is a constrained but unregularized iterative image restoration algorithm. The ICTM and Carrington algorithms, however, incorporate Tikhonov regularization to suppress undesired solutions. Conchello has derived an algorithm that incorporates Tikhonov regularization into the Richardson–Lucy algorithm (Conchello & McNally, 1996). Incorporation of Tikhonov regularization yields

$$\hat{\mathbf{f}}^{k+1}_{regularized} = \frac{-1 + \sqrt{1 + 2\lambda \hat{\mathbf{f}}^{k+1}}}{\lambda} \tag{21}$$

with $\hat{\mathbf{f}}^{k+1}$ given by Eq. (20). Using l'Hôpital's rule it is easy to show that (21) becomes (20) when $\lambda \to 0$. We will refer to this algorithm as the RL–Conchello algorithm.

3. Regularization parameter

In this section we introduce various methods that have been proposed to estimate the regularization parameter for the linear Tikhonov–Miller filter. In the following section we will test whether these methods are suitable to be applied to non-linear restoration algorithms for the restoration of Poisson distorted images.

3.1. The SNR method

The signal-to-noise ratio (SNR) method sets λ equal to the inverse of the signal-to-noise ratio

$$\lambda_{SNR} = \frac{\epsilon}{E} = \frac{\epsilon}{\sum |\mathbf{g} - \mathbf{b}|^2} \tag{22}$$

with ϵ the total power of the noise and E the total power of the object in the image.

3.2. The method of constrained least squares

In Galatsanos & Katsaggelos (1992) the methods of constrained least squares, generalized cross-validation and maximum likelihood are described to determine λ for the Tikhonov–Miller algorithm. These methods define different criteria to balance the difference between the recorded data and the reblurred restoration result, given the amount of noise present in the image.

The method of constrained least squares (CLS) finds a λ_{CLS} such that the mean-square-error between the recorded data and the blurred restored data equals the noise power,

$$\sum \left|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}(\lambda)\right|^2 = \sum \left|(\mathbf{I} - \mathbf{HA}(\lambda))\mathbf{g}\right|^2 = \epsilon \tag{23}$$

with $\mathbf{I}$ denoting the unity matrix and ϵ the total noise power. $\mathbf{A}$ is the linear Tikhonov–Miller filter,

$$\mathbf{A}(\lambda) = (\mathbf{H}^T\mathbf{H} + \lambda \mathbf{C}^T\mathbf{C})^{-1}\mathbf{H}^T \tag{24}$$

We determine λ_{CLS} numerically using Brent's method (Press *et al.*, 1992) to find the zero crossing of the function

$$\sum \left|(\mathbf{I} - \mathbf{HA}(\lambda))\mathbf{g}\right|^2 - \epsilon \tag{25}$$

3.3. The method of generalized cross-validation

The method of generalized cross-validation (GCV) is derived from the leave-one-out principle (Galatsanos & Katsaggelos, 1992). For every pixel its associated Tikhonov–Miller filter is calculated using all but the pixel under consideration. The cross-validation function is derived from the mean-square-error between the original data and the result after filtering each pixel with its associated Tikhonov–Miller filter. The generalized cross-validation function (CV),

$$CV(\lambda) = \frac{\sum |(\mathbf{I} - \mathbf{HA}(\lambda))g|^2}{[\mathrm{trace}(\mathbf{I} - \mathbf{HA}(\lambda))]^2} \tag{26}$$

can be efficiently evaluated in the discrete Fourier domain

(Galatsanos & Katsaggelos, 1992)[2],

$$CV(\lambda) = \frac{\sum_{\omega} \frac{\lambda^2 |C(\omega)|^4 |G(\omega)|^2}{(|H(\omega)|^2 + \lambda |C(\omega)|^2)^2}}{\left(\sum_{\omega} \frac{\lambda |C(\omega)|^2}{|H(\omega)|^2 + \lambda |C(\omega)|^2}\right)^2} \tag{27}$$

The minimum of this function can be found without prior knowledge of the noise variance (Reeves & Mersereau, 1992). We have determined the minimum of the cross-validation function numerically with Brent's one-dimensional search algorithm (Press *et al.*, 1992).

3.4. The method of maximum likelihood

An alternative method to determine λ without prior knowledge of the noise variance has been named the maximum likelihood method (ML) by (Galatsanos & Katsaggelos, 1992). It is based on a stochastic approach, which assumes $\sqrt{\lambda}\mathbf{Cf}$ and the noise to be Gaussian distributed. The former assumption can be met with a proper choice of the regularization matrix $\mathbf{C}$ (Galatsanos & Katsaggelos, 1992). The derived maximum likelihood function,

$$ML(\lambda) = \frac{\mathbf{g}^T(\mathbf{I} - \mathbf{HA}(\lambda))\mathbf{g}}{(\det[\mathbf{I} - \mathbf{HA}(\lambda)])^{1/M}} \tag{28}$$

can be evaluated in the discrete Fourier domain (Galatsanos & Katsaggelos, 1992)

$$ML(\lambda) = \frac{\sum_{\omega} \frac{\lambda |C(\omega)|^2 |G(\omega)|^2}{|H(\omega)|^2 + \lambda |C(\omega)|^2}}{\left(\prod_{\omega} \frac{\lambda |C(\omega)|^2}{|H(\omega)|^2 + \lambda |C(\omega)|^2}\right)^{1/\Omega}} \tag{29}$$

with Ω the number of Fourier coefficients. We used Brent's one-dimensional search algorithm to find the minimum of (29).

3.5. The Golden search method

We compare these four methods for determining λ (SNR, CLS, GCV and ML) with the 'optimal' value for λ. We define the optimal value of λ as the value for which the image restoration algorithm (either the ICTM algorithm or Carrington's algorithm) produces the smallest mean-square-error. We have used the Golden search algorithm to find the optimal value numerically (Press *et al.*, 1992).

3.6. Estimation of the noise variance

Both the CLS method and the SNR method use the total noise power ϵ. Therefore, they require prior knowledge about the noise variance. Often this knowledge is not available. It is possible however, to estimate the noise variance using the GCV method (Galatsanos & Katsaggelos, 1992). Denoting the value of the regularization parameter determined by the GCV method with λ_{GCV}, the noise variance can be estimated by

$$\hat{\sigma}^2 = \frac{|(\mathbf{I} - \mathbf{HA}(\lambda_{GCV}))\mathbf{g}|^2}{\mathrm{trace}[\mathbf{I} - \mathbf{HA}(\lambda_{GCV})]} \tag{30}$$

Although this estimation of the noise variance assumes that the data are distorted by additive Gaussian noise, the results presented in this section show that Eq. (30) can also be used to estimate the average noise variance in a fluorescence microscopical image distorted by Poisson noise. Using the experimental conditions as described in section 4 we performed a simulation experiment in which this method for estimation of the noise variance was tested for confocal microscopical imaging.

Figure 2 shows the relative error (the estimated value minus the true value divided by the true value) of the estimated noise variance in the top-left graph. The top-right graph shows the relative error in the regularization value determined with the CLS method using the estimated and true variance value. We define the true value of the variance as the mean variance over the image, which in the case of Poisson noise is equal to the mean intensity of the acquired image. The bottom graph shows the relative difference between the mean-square-error performance of the ICTM algorithm with the regularization parameter determined with the CLS method using an estimated variance and the true value of the variance. From Fig. 2 we conclude that using the estimated variance produces values of the regularization parameter very close to those found by the CLS method using the true value for the variance. Furthermore, it is shown that the mean-square-error performance of the ICTM algorithm is hardly influenced by using the estimated variance.

The values of the regularization parameter as determined by SNR, CLS, GCV and ML are plotted in Fig. 3 as a function of the signal-to-noise ratio. We define the optimal value of the regularization parameter as the value that minimizes the mean-square-error between the original image $\mathbf{f}$ and the restored image $\hat{\mathbf{f}}$. We have determined the optimal values of the regularization parameter with a Golden search algorithm (Press *et al.*, 1992) for both the ICTM algorithm and the Carrington algorithm. These values are also plotted in Fig. 3.

4. Comparison of the regularization parameter methods

The simulation experiments described in this section are performed on images with a size of $128 \times 128 \times 64$ pixels. We used a two times oversampled confocal point spread function,

[2] See Golub *et al.* (1979) for a step-by-step derivation of the generalized cross validation function.

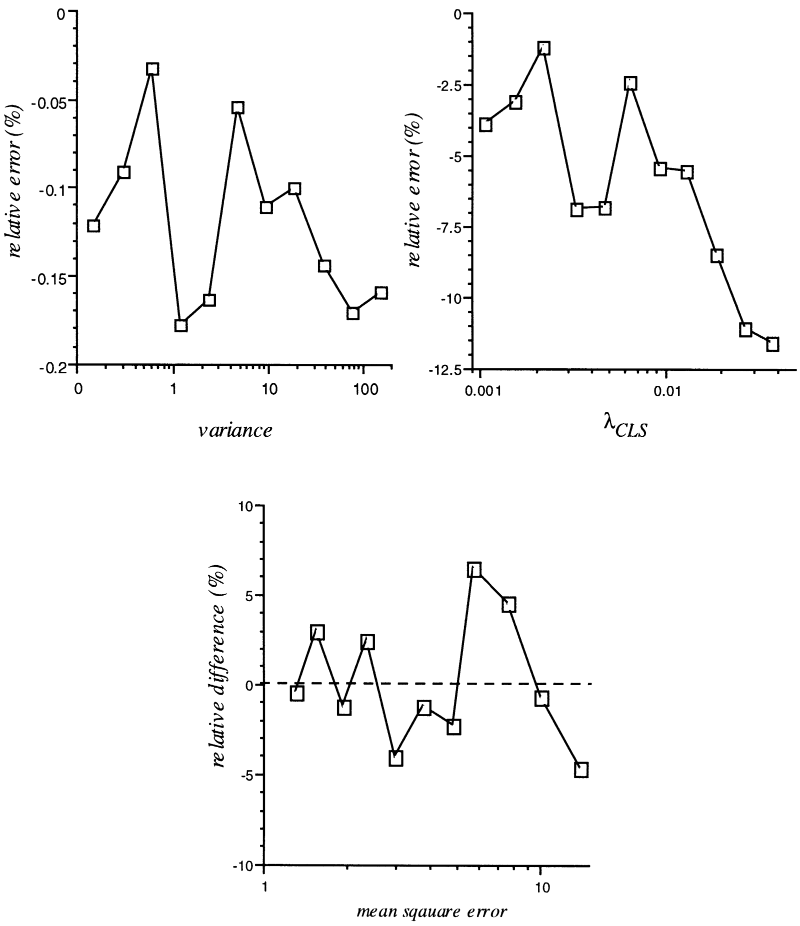


Fig. 2. The top left graph shows the relative error of the estimated variance using GCV. The top right graph shows the relative error of the regularization parameter determined with CLS using the variance estimated by GCV. The bottom graph shows the relative difference in mean-square-error of ICTM with CLS regularization due to estimation of the variance.

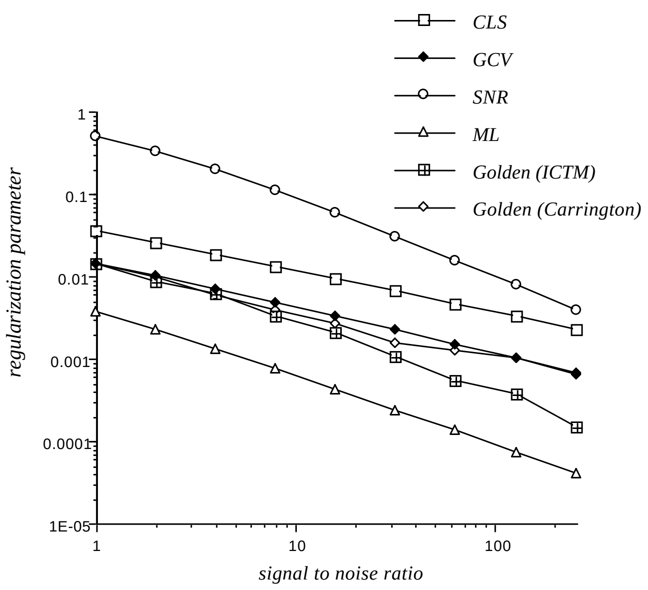


Fig. 3. The value of the regularization parameter, using SNR, CLS, GCV and ML, as a function of the signal-to-noise ratio. The optimal values (Golden) of λ for the ICTM and Carrington's algorithm are also plotted.

an NA of 1.3, an excitation wavelength of 480 nm, an emission wavelength of 530 nm[3], a refractive index of 1.515 and a backprojected pinhole size of 300 nm. Spheres with a diameter of 1000 nm were used as objects. The sampling distance was 25.0 nm lateral and 85.0 nm axial. We varied the signal-to-noise ratio (as defined by the ratio of the object energy and the noise energy) from 1.0 to 256.0 (0.0 dB to 24.1 dB), which corresponds to a conversion factor ranging from 0.14 to 35.8, with an object intensity of 200.0 ADU and a background of 20.0 ADU.

We have based our implementation of the various methods for determining the regularization parameter on sources provided by Peter J. Verveer. We reimplemented these algorithms in the scientific image processing library DIPlib (www.ph.tn.tudelft.nl/DIPlib, Pattern Recognition Group, Delft University of Technology, The Netherlands) to incorporate the methods to estimate the signal-to-noise ratio as presented in section 3.6.

Figure 4 shows the mean-square-error performances of the ICTM and Carrington algorithms for the various

[3] These excitation and emission wavelength correspond to the excitation and emission spectra of popular fluophores like FITC and Nile-Red.

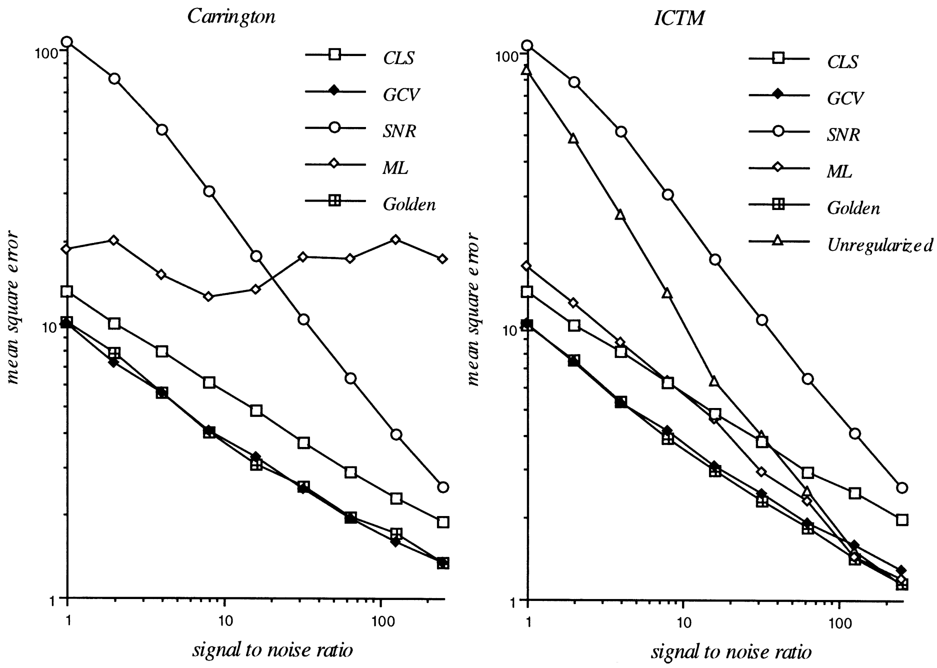


Fig. 4. The value of the mean-square-error of the restoration result of the Carrington (left) and ICTM (right) algorithm. The regularization parameter was determined with the CLS, GCV, SNR and ML method and a Golden search. The performance of an unregularized ICTM is also shown.

methods used to estimate the regularization parameter. We have also included the performance of the ICTM with no regularization (λ set to zero). Owing to numerical instability (see also the low performance of the Carrington algorithm for low values of λ as estimated with ML), we were not able to include these results for the Carrington algorithm. All the results shown in Fig. 4 are averages over eight experiments. All data points have a coefficient-of-variation of less than 5%.

Figure 4 clearly shows that high values of the regularization parameter as determined by the SNR method result in a low performance caused by an oversmoothed restoration result (see Fig. 5).

The CLS method is also known to overestimate the value of the regularization parameter (Galatsanos & Katsaggelos, 1992). Therefore, the CLS method will produce somewhat smooth (restored) images resulting in a suboptimal performance.

The Tikhonov functional (see Eq. (7)) includes a regularization matrix **R**, which is supported by the ICTM algorithm. The Carrington algorithm does not support this regularization matrix. We have therefore chosen to make **R** equal to the unity matrix **I** and not to use a specific regularization matrix **R** that would make **Rf** Gaussian distributed, as required by the ML algorithm. This explains the weak performance of the ML method in combination with the ICTM algorithm. In our experiments the ML method produces very small values for the regularization parameter, which could not be handled properly by the Carrington algorithm. We attribute the low performance of the Carrington algorithm in this case to numerical instability of the algorithm (we believe that this is related to the 'extra' transform (Eq. (12)) that the Carrington algorithm incorporates).

The performances of the GCV and the Golden search method are equal within the accuracy of the experiment. Clearly the GCV method is our method of choice for estimating the regularization parameter.

Figure 5 shows centre x-y and x-z slices of the restoration results obtained with the ICTM and Carrington algorithms using the different methods to determine the regularization parameter. At first, the result obtained with the SNR method might seem the most appealing since it lacks the structures or texture present in the other results that are not present in the original object. These structures are caused by the noise realization in the acquired image. The SNR method imposed the strongest regularization at the SNR chosen in this figure (see Fig. 4), which smooths not only the noise-induced texture but also the restoration result.

5. The regularization parameter of the RL–Conchello algorithm

In the previous sections we have shown that the GCV method is a good method for determining the regularization

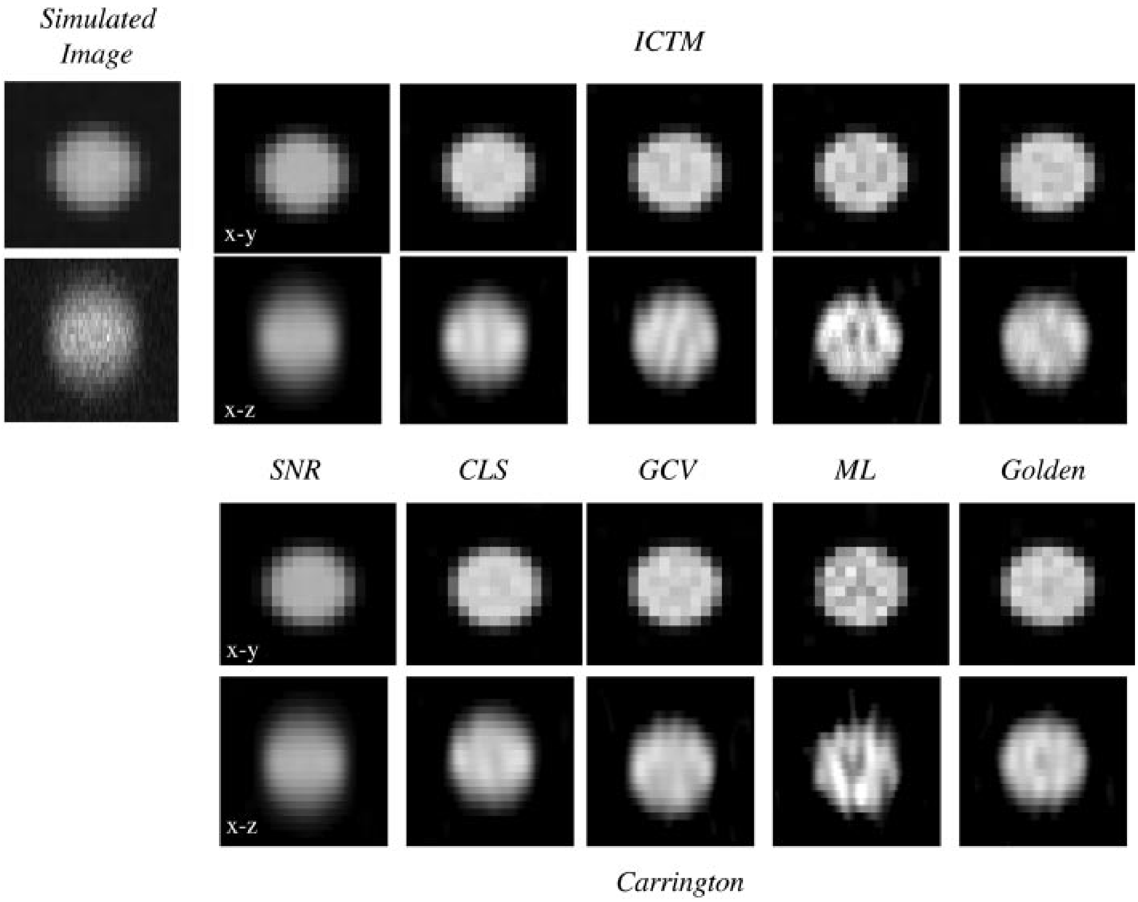


Fig. 5. Restoration results of the ICTM and Carrington algorithms with the regularization parameter determined by the SNR, CLS, GCV, ML, and Golden search method. The pictures show the centre *x-y* and *x-z* slices of the centre (80×80 pixels *x-y*, 80×40 pixels *x-z*) of an image of a sphere with a 0.50-μm radius and a signal-to-noise ration of 8.0.

parameter for the ICTM algorithm and the Carrington algorithm. The method is computationally feasible as we use the Tikhonov–Miller algorithm to compute an estimate for **f**.

The RL–Conchello algorithm is based on a regularized maximum likelihood function for images distorted by Poisson noise. No linear solution exists for this functional. We have therefore approximated the functional such that it can be solved linearly. This linear solution can then be used to estimate the regularization parameter with a feasible computational complexity.

The RL–Conchello maximizes the regularized loglikelihood functional of a translated Poisson process

$$L(\hat{\mathbf{f}}) = -\sum \mathbf{H}\hat{\mathbf{f}} + \mathbf{m}^T \ln(\mathbf{H}\hat{\mathbf{f}} + \mathbf{b}) - \lambda\|\hat{\mathbf{f}}\|^2 \tag{31}$$

Maximizing Eq. (31) is equal to minimizing

$$c + \sum \mathbf{H}\hat{\mathbf{f}} - \mathbf{m}^T \ln(\mathbf{H}\hat{\mathbf{f}} + \mathbf{b}) + \lambda\|\hat{\mathbf{f}}\|^2 \tag{32}$$

where c is a scalar constant. If we choose c to be the log likelihood of the recorded data, Eq. (32) becomes

$$\sum \mathbf{g} - \mathbf{m}^T \ln \mathbf{m} + \sum \mathbf{H}\hat{\mathbf{f}} - \mathbf{m}^T \ln(\mathbf{H}\hat{\mathbf{f}} + \mathbf{b}) + \lambda\|\hat{\mathbf{f}}\|^2 \tag{33}$$

A second order Taylor series expansion of ??? around **g** yields

$$\sum(2\mathbf{m})^{-1}(\mathbf{H}\hat{\mathbf{f}} - \mathbf{g})^2 + \lambda\|\hat{\mathbf{f}}\|^2 = \sum(\mathbf{H}\hat{\mathbf{f}} + \mathbf{g})^2 + 2\lambda \sum \mathbf{m}\hat{\mathbf{f}}^{\,2} \tag{34}$$

To be able to find a linear solution of Eq. (34) in the form of a Tikhonov–Miller filter, we need to approximate **m** by a constant. Using the mean of **m** Eq. (34) becomes

$$\sum(\mathbf{H}\hat{\mathbf{f}} - \mathbf{g})^2 + 2\lambda\bar{m}\|\hat{\mathbf{f}}\|^2 \tag{35}$$

Using one of the methods to determine the regularization parameter for the Tikhonov–Miller filter, we can determine the regularization parameter for the RL–Conchello algorithm

$$\lambda_{RL-Conchello} = \lambda_{TM}/2\bar{m} \tag{36}$$

Thompson (1989) has used similar approximations for determining the regularization parameter for a mean-square-error functional regularized by an entropy measure.

Figure 6 shows the values of the regularization value for the RL–Conchello algorithm determined by the CLS, GCV, SNR and ML algorithms. Again we used the Golden search algorithm to find the optimal value of the regularization parameter, which is also included.

The mean-square-error performance of the RL–Conchello algorithm is shown in Fig. 7 with the regularization parameter determined by the five different methods as a function of the signal-to-noise ratio. We have also included the performance of the Richardson–Lucy algorithm, the unregularized version of the RL–Conchello algorithm.

It is clear from Fig. 6 that the optimal value of the regularization value is determined considerably less accurately by any of the four proposed methods than in case of the ICTM and Carrington algorithms. One reason for this could be the approximation we had to make in the derivation of Eq. (35), where we approximated **m** by $\bar{\mathbf{m}}$. However, the GCV method still produces results that are considerably better in the mean-square-error sense than the results of the unregularized Richardson–Lucy algorithm.

6. First estimates

The investigated iterative restoration algorithms need a first estimate to start their iterations. In this section we compare the performance of the ICTM, Carrington and Richardson–Lucy algorithms using different first estimates. We have tested the following first estimates:

- *the recorded image*

An obvious choice for the first estimate is to use the image **m** acquired by the microscope. For the ICTM and Carrington algorithms we have subtracted the (constant) background **b** from **m** to make the summed intensity of ($\mathbf{m} - \mathbf{b}$) equal to that of the original object. We have defined $\mathbf{m} - \mathbf{b}$ as **g** (see Eq. (2)). For the RL–Conchello and Richardson–Lucy algorithms, we used **m** to guarantee the positivity of the first estimate.

- *the mean of the recorded image*

Kaufman (1987) has argued that starting from a uniform first guess is a good approach if no other reasonable guess as a first estimate can be obtained. We have used the mean of **g** for the ICTM and Carrington algorithms and the mean of **m** for the RL–Conchello and Richardson–Lucy algorithms as a uniform first estimate.

- *a smoothed version of the recorded image*

The recorded image contains the smoothed object as well as noise. To prevent the restoration algorithm adapting to noise realizations in the recorded image one could reduce the noise in **g** (**m** for the RL–Conchello and Richardson–Lucy algorithms) by means of local smoothing. We have smoothed **g** with a Gaussian filter with a sigma of 2.0 pixels (this imposes a smoothing comparable to the smoothing of the point spread function used in this experiment).

- *a homogeneous noise image*

The mean of **g** (or **m**) is a constant image in the spatial domain and an impulse in the Fourier domain. In that case the restoration algorithms have to restore the complete Fourier domain starting from an impulse. It is interesting to see if it is beneficial to start with a first estimate that fills the Fourier domain. We have generated such an image by generating a constant image distorted by Poisson noise. The mean of this image has been set to the mean of **g**.

- *the result of Tikhonov-Miller restoration*

The iterative algorithms based on the Tikhonov functional should converge more quickly to their solutions if a first estimate close to the final solution is provided. We used the restoration result produced by the linear Tikhonov–Miller filter to test this hypothesis. For the RL–Conchello and Richardson–Lucy algorithms we have clipped this result at zero to guarantee the positivity of the first estimate.

- *the original object*

Ideally, the restoration algorithms should produce the original object. Using the original object as a first estimate enables us to test whether this is the solution found by these algorithms under realistic circumstances.

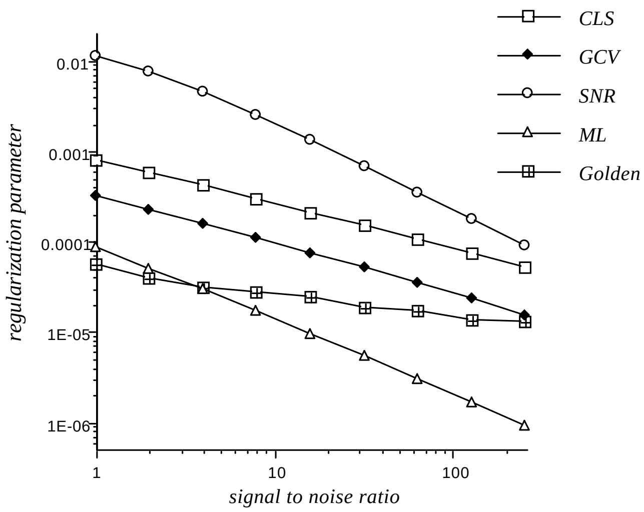


Fig. 6. The value of the regularization parameter for the RL–Conchello algorithm determined with SNR, CLS, GCV and ML, as a function of the signal-to-noise ratio. The optimal values (Golden) are also plotted.

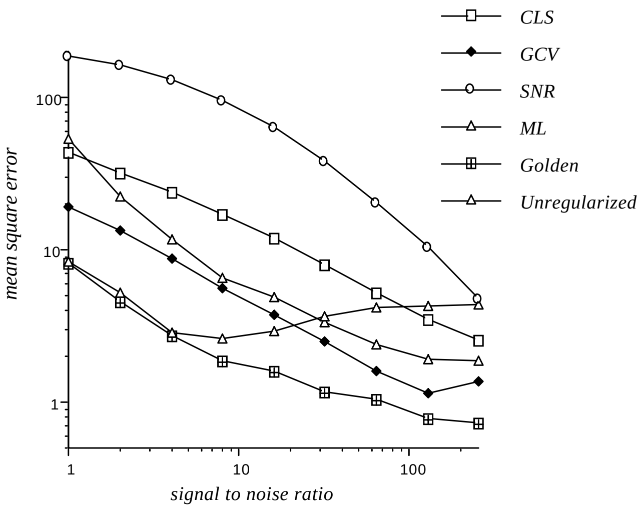


Fig. 7. The value of the mean-square-error of the restoration result of the RL–Conchello algorithm. The regularization parameter is determined with the CLS, GCV, SNR and ML method and a Golden search. The performance of the unregularized Richardson–Lucy algorithm is also shown.

- *a zero image for the Carrington algorithm only*

The Carrington algorithm finds the estimate of **f** by optimizing a functional as function of a transformed parameter. Therefore, none of the first estimates proposed above is a really good estimate for the Carrington algorithm. Carrington has proposed the use of an image set to zero as the first estimate (Carrington, 1990).

We have measured the mean-square-error and the I-divergence of the ICTM algorithm, the Carrington algorithm, the RL–Conchello algorithm and the Richardson–Lucy algorithm as a function of the number of iterations for the various first estimates. The measured mean-square-error and I-divergence values of the first 200 iterations are shown in Fig. 8 for the ICTM algorithm, in Fig. 9 for the Carrington algorithm, in Fig. 10 for the RL–Conchello algorithm, and in Fig. 11 for the Richardson–Lucy algorithm. The data show the averages over eight repetitions of the experiment using different noise realizations. Note that we have used different scales in these graphs to account for the differences in the range of the data.

From these figures we conclude that the ICTM and RL–Conchello algorithms using the various first estimates converge to the same end result in both mean-square-error and I-divergence sense. Using the relatively noisy first estimates (the recorded image and homogeneous noise image) the Carrington and Richardson–Lucy algorithms produce a result with a somewhat poorer performance. Also the zero image first estimate for the Carrington algorithm produces a slightly poorer performance.

The ICTM algorithm shows the fastest convergence of the four algorithms tested. The results also show that none of the first estimates is a really good first estimate for the Carrington algorithm. The convergence is slow and rough when compared to the other results.

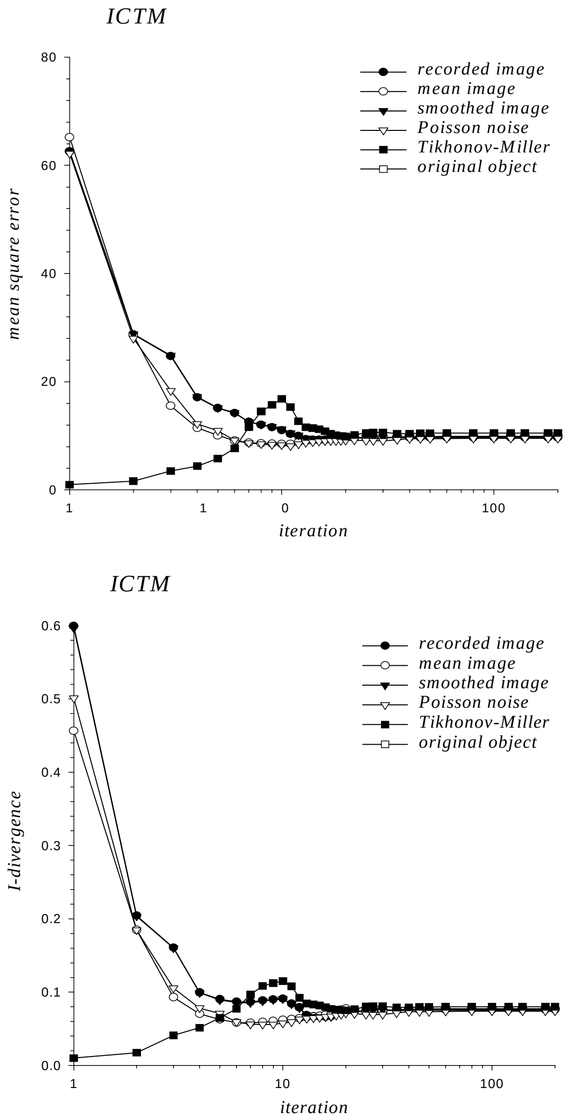


Fig. 8. The mean-square error and I-divergence for different first estimates as a function of the number of iterations performed by the ICTM algorithm.

The three regularized algorithms (ICTM, Carrington and RL–Conchello) converge to a result that is very similar in both the mean-square-error and the I-divergence sense. The performance of the (unregularized) Richardson–Lucy algorithm is lower than that of the other three and does not converge to a minimum in both mean-square-error and I-divergence sense. The performance even decreases somewhat for a large number of iterations. This can be explained by the lack of regularization in the Richardson–Lucy algorithm, which allow the restoration result to adapt to the noise present in the acquired image (Conchello *et al.*, 1994).

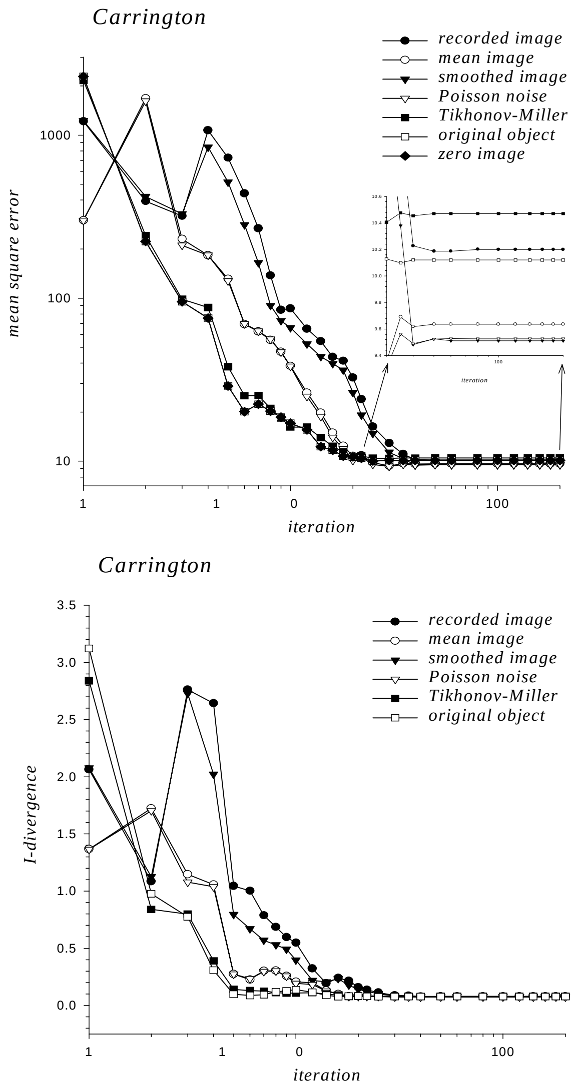


Fig. 9. The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the Carrington algorithm.

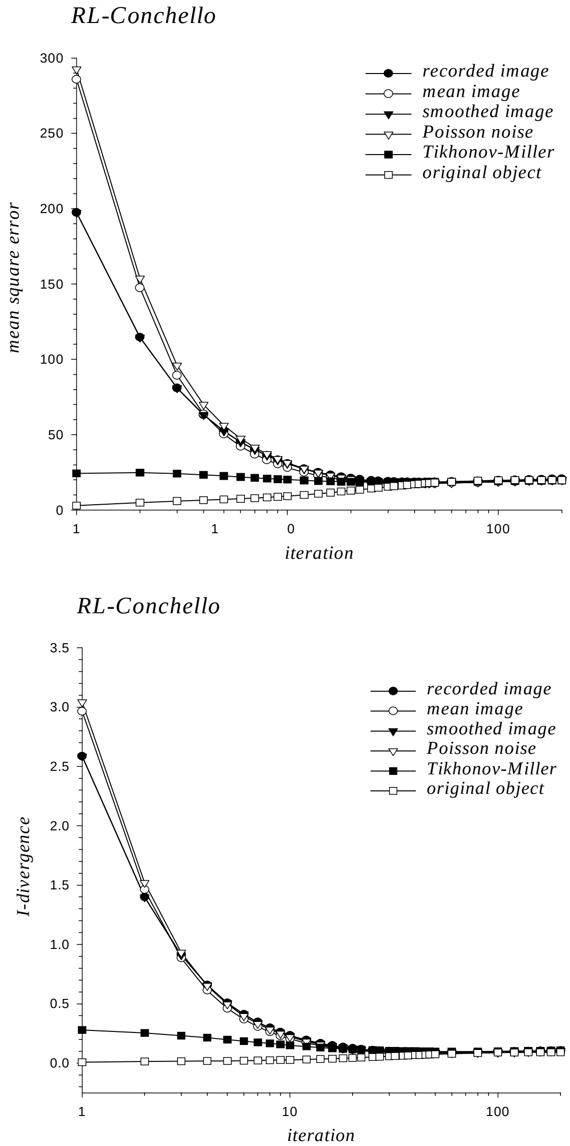
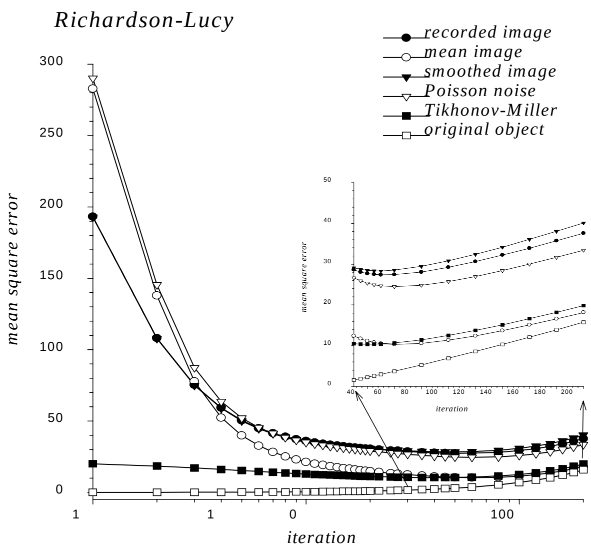


Fig. 10. The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the RL–Conchello algorithm.

Only the ICTM and Carrington algorithms benefit from using the smoothed image as a first estimate. The ICTM algorithm converges quickly for this first estimate, whereas the Carrington algorithm produces a slightly better restoration result in mean-square-error sense. The convergence of this first estimate is worse than that of the recorded image for both the RL–Conchello and the Richardson–Lucy algorithm. Using the restoration result of the Tikhonov–Miller filter as a first estimate gives by far the fastest convergence for all algorithms except for the Carrington algorithm.

Even when the original object is used as a first estimate, the various algorithms converge to the same solution in the mean-square-error sense as obtained with the other first estimates. This can be explained by the regularization. The Tikhonov regularization can be regarded as a Gaussian a priori model of the original (Andrews & Hunt, 1977; Verveer & Jovin, 1997). Because the intensities of our original object are not Gaussian distributed, the algorithms will find a solution, which is a balance between the distance to the original object and the Gaussian model imposed by the regularization.

Although the result was a surprise, we have not investigated why the ICTM algorithm produces a maximum mean-square-error at about 10 iterations when the original object is used as a first estimate. This phenomenon could be related to the type of object or the type of microscope (confocal imaging) used in the simulations.



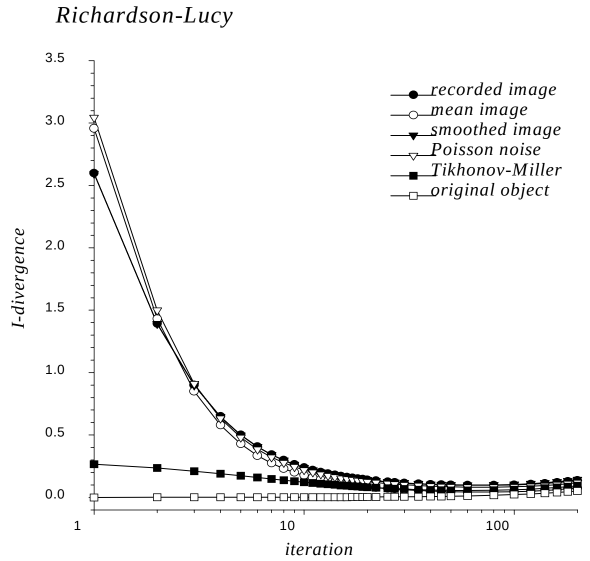


Fig. 11. The mean-square-error and I-divergence for different first estimates as a function of the number of iterations performed by the Richardson–Lucy algorithm.

7. Conclusions

The performance of non-linear iterative image restoration algorithms is dependent on various parameters. In this paper we have investigated the impact of the regularization parameter and the choice of first estimate.

We have compared various methods for determining the regularization parameter, which has been developed by Galatsanos (Galatsanos & Katsaggelos, 1992) for linear Tikhonov restoration of images distorted with additive Gaussian noise. We have tested the ability of these algorithms to determine the regularization parameter of non-linear restoration algorithms for the restoration of images distorted with Poisson noise. Our results show that for the ICTM and Carrington algorithms the method of generalized cross validation (GCV) provides an efficient way of producing a value of the regularization parameter which is close to its optimal value. We defined the optimal value to be that value that produced the minimum mean-square-error between the restoration result and the original object. We determined this optimal value using the Golden search algorithm (Press *et al.*, 1992), a numerical minimization algorithm. Over a large range of the signal-to-noise ratio, the GCV method produced practically the same value for the regularization parameter as was obtained with this search algorithm. (This 'optimal' procedure can only be performed on synthetic data as it requires the non-blurred data as well.) This indicates that the assumption of additive Gaussian noise and the use of the linear Tikhonov restoration filter by the GCV method does not influence its performance even when used on images distorted with Poisson noise at a low signal-to-noise ratio.

We have also derived a method for estimating the regularization parameter for the RL–Conchello algorithm, a Tikhonov regularized version of the Richardson–Lucy algorithm. For this algorithm, the GCV method is again the best method, among the tested methods, for estimating the regularization parameter. However, the values obtained with the GCV method are not as close to the 'optimal' value (again obtained with a Golden search), as it was for the ICTM and Carrington algorithms. This is probably caused by the approximation of the acquired image by its mean value, which was necessary in the derivation of the regularization parameter estimation for this algorithm.

We have investigated the influence of various choices of a first estimate as used by the iterative image restoration algorithms investigated here. We found that most algorithms converged to the same solution for the tested choices. Only the Richardson–Lucy algorithm failed to converge to a steady solution but showed a decrease in performance for a large number of iterations. This can be explained by the unregularized nature of this algorithm, making it sensitive to the noise realization present in the acquired image.

Acknowledgements

This work was partially supported by the Royal Netherlands Academy of Arts and Sciences (KNAW) and by the Rolling Grants program of the Foundation for Fundamental Research in Matter (FOM).

References

Andrews, H.C. & Hunt, B.R. (1977) *Digital Image Restoration*. Prentice Hall, Englewood Cliffs, NJ.

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

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.

Conchello, J.A., Kim, J.J. & Hansen, E.W. (1994) Enhanced 3-dimensional reconstruction from confocal scanning microscope images. 2. Depth discrimination versus signal-to-noise ratio in partially confocal images. *Appl. Optics*, **33**, 3740–3750.

Conchello, J.-A. & McNally, J.G. (1996) Fast regularization technique for expectation maximization algorithm for optical sectioning microscopy. *Three-Dimensional Microscopy: Image Acquisition and Processing III* Vol. 2655 (ed. by C. J. Cogswell, G. S. Kino and T. Wilson), pp. 199–208. SPIE, San Jose.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. 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.

Golub, G.H., Heath, M. & Wahba, G. (1979) Generalized cross validation as a method for choosing a good Rigde parameter. *Technometrics*, **21**, 215–223.

Holmes, T.J. (1988) Maximum-likelihood image restoration adapted for noncoherent optical imaging. *J. Opt Soc. Am. A*, **5**, 666–673.

Kaufman, L. (1987) Implementing and accelerating the EM algorithm for positron emission tomography. *IEEE Trans. Med. Imaging*, **MI-6**, 37–51.

van Kempen, G.M.P. (1999) *Image Restoration in Fluorescence Microscopy.* PhD Thesis, Delft University of Technology, Delft, The Netherlands.

van Kempen, G.M.P., van Vliet, L.J. & Verveer, P.J. (1997) Application of image restoration methods for confocal fluorescence microscopy. *3-D Microscopy: Image Acquisition and Processing IV Proceedings of the SPIE.* Vol. 2984 (ed. by J.-A. Conchello, C. J. Cogswell and T. Wilson), pp. 114–124. SPIE, San Jose.

van Kempen, G.M.P., van Vliet, L.J., Verveer, P.J., van der Voort, H.T.M. (1997) A quantitative comparison of image restoration methods for confocal microscopy. *J. Micros*. **185**, 354–365.

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

Miller, K. (1970) Least squares methods for ill-posed problems with a prescribed bound. *SIAM J. Math Anal.* **1**, 52–74.

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.

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

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.

Thompson, A.M. (1989) On the use of quadratic regularisation within maximum entropy image restoration. *Maximum Entropy and Bayesian Methods* (ed. by J. Skilling), pp. 497–504. Kluwer Academic Publishers, Dordrecht.

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.

Verveer, P.J., Gemkow, M.J. & Jovin, T.M. (1999) A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy. *J. Microsc.* **193**, 50–61.

Verveer, P.J. & Jovin, T.M. (1997) Efficient super-resolution restoration algorithms using MAP estimations with application to fluorescence microscopy. *J. Opt Soc. Am. A*, **14**, 1696–1706.

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