

Efficient superresolution restoration algorithms using maximum *a posteriori* estimations with application to fluorescence microscopy

Peter J. Verveer and Thomas M. Jovin

Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Postfach 2841, D-37018 Göttingen, Germany

Received July 22, 1996; revised manuscript received December 12, 1996; accepted January 27, 1997

We present efficient algorithms for image restoration by using the maximum *a posteriori* (MAP) method. Assuming Gaussian or Poisson statistics for the noise and either a Gaussian or an entropy prior distribution for the image, corresponding functionals are formulated and minimized to produce MAP estimations. Efficient algorithms are presented for finding the minimum of these functionals in the presence of nonnegativity and support constraints. Performance was tested by using simulated three-dimensional (3-D) imaging with a fluorescence confocal laser scanning microscope. Results are compared with those from two existing algorithms for superresolution in fluorescence imaging. An example is given of the restoration of a 3-D confocal image of a biological specimen. © 1997 Optical Society of America [S0740-3232(97)01708-0]

1. INTRODUCTION

It has been recognized for a long time that image restoration techniques can enhance considerably the resolution of images obtained by microscopy. An extensive literature exists on image restoration techniques applied in numerous fields including astronomy and microscopy. Our work was motivated by the need for processing three-dimensional (3-D) fluorescent images acquired by conventional and confocal microscopy. However, the algorithms presented here are of general applicability.

The most successful algorithms that have been applied in fluorescence microscopy are nonlinear in nature and apply prior knowledge regarding the objects to be restored and the noise in the image. Prominent examples are the constrained iterative algorithm of Agard *et al.*,¹ the expectation-maximization (EM) algorithm,²⁻⁴ the regularized algorithm of Carrington,^{5,6} and the iterative constrained Tikhonov–Miller (ICTM) algorithm.^{7,8} These algorithms possess superresolution properties, allowing partial recovery of frequencies that are lost in the imaging process. The observation that superresolution performance can be achieved by using prior knowledge of the objects has recently raised considerable interest in image restoration.^{9,10}

In this paper we use a maximum *a posteriori* (MAP) approach to solve the restoration problem, deriving functionals that are minimized to produce MAP estimates for the objects. We propose an efficient iterative approach for minimizing these functionals, subject to nonnegativity and support constraints. We discuss the relation of our work to existing approaches. Simulated 3-D confocal images have been used to analyze the performance of the proposed algorithms and to compare them with those of Holmes and of Carrington. We also present an example of the restoration of a biological sample imaged with a confocal fluorescence microscope.

2. DERIVATION OF THE ALGORITHMS

A. Imaging Model

We assume that we can model the imaging process by a linear system of equations, which for the 3-D case after sampling is given by

$$g(x, y, z) = N \left[\sum_{i,j,k} h(x, y, z, i, j, k) f(i, j, k) + b(x, y, z) \right], \quad (1)$$

where $g(x, y, z)$ represents the 3-D blurred image, $f(i, j, k)$ represents the original 3-D object, $h(x, y, z, i, j, k)$ represents the blurring function, and $b(x, y, z)$ represents a distinct known background in the image. The function $N(\cdot)$ denotes the application of a noise process to the image, for which we consider Gaussian or Poisson distributions.

In this paper a matrix notation is used. The image is converted into a vector by concatenating its rows, leading to the imaging model

$$\mathbf{g} = N(\mathbf{H}\mathbf{f} + \mathbf{b}), \quad (2)$$

where the vectors \mathbf{f} and \mathbf{g} of length M denote the original object and the blurred image, respectively. The $M \times M$ matrix \mathbf{H} is the blurring matrix. Note that the images need not be three dimensional but can have any dimension.

Solving \mathbf{f} from Eq. (2) is an ill-conditioned problem, since \mathbf{H}^{-1} may (in general) not exist. If \mathbf{H} denotes a space-invariant process, then the matrix product can be evaluated as a convolution with a point-spread function (PSF), which is equivalent to multiplication in the Fourier domain. Because of diffraction effects, the Fourier transform of the PSF of an optical microscope falls off rapidly and vanishes completely after some cutoff frequency. In

such a case \mathbf{H}^{-1} does not exist. The pseudoinverse solution of Eq. (2) does exist but is unstable in the presence of noise. These problems have led to the application of regularization techniques and prior knowledge to obtain a stable estimate of \mathbf{f} .

B. Maximum a Posteriori Image Restoration Approach

Define a prior density $p(\mathbf{f})$ for the distribution of the intensity of the object and a conditional density $p(\mathbf{g}|\mathbf{f})$ for the measured image. The posterior density $p(\mathbf{f}|\mathbf{g})$ is given by Bayes's rule:

$$p(\mathbf{f}|\mathbf{g}) = p(\mathbf{g}|\mathbf{f})p(\mathbf{f})/p(\mathbf{g}). \quad (3)$$

The MAP solution for \mathbf{f} is obtained by maximizing the right-hand-side of Eq. (3), which is equivalent to maximizing the numerator. If a uniform distribution is chosen for $p(\mathbf{f})$, this reduces to the maximum likelihood (ML) solution. One possible choice for the prior density is the Gaussian distribution

$$p(\mathbf{f}) = C \exp\left[-\frac{1}{2\tau^2} \|\mathbf{C}(\mathbf{f} - \mathbf{m})\|^2\right], \quad (4)$$

where $\|\cdot\|^2$ is the Euclidean norm. In image restoration \mathbf{C} and \mathbf{m} are commonly referred to as the regularization matrix and the model, respectively. Their choice is discussed in Subsection 2.E. A Gaussian prior may not apply in general, but we will show that this choice of $p(\mathbf{f})$ leads to a practical way for regularizing the solution.

It has been argued by Skilling¹¹ that in the absence of further prior knowledge, the entropy distribution is the only consistent prior distribution for positive, additive images. Therefore we use the entropy prior:

$$p(\mathbf{f}) = C \exp[\rho S(|\mathbf{C}\mathbf{f}|, |\mathbf{C}\mathbf{m}|)], \quad (5)$$

where $S(\mathbf{x}, \mathbf{y})$ is the entropy of \mathbf{x} relative to a model \mathbf{y} , given by

$$S(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^M \left[\mathbf{x}_i - \mathbf{y}_i - \mathbf{x}_i \ln\left(\frac{\mathbf{x}_i}{\mathbf{y}_i}\right) \right]. \quad (6)$$

The model \mathbf{y} represents prior knowledge of \mathbf{x} , since it determines the global maximum of S and therefore the most likely value of \mathbf{x} according to the entropy distribution. Thus a particular intensity distribution of \mathbf{x} can be made more likely by adjusting the intensity distribution of \mathbf{y} accordingly. In the absence of any prior knowledge, \mathbf{y} is chosen to have a constant intensity. Note that in Eq. (5) we substituted $\mathbf{x} = |\mathbf{C}\mathbf{f}|$ and $\mathbf{y} = |\mathbf{C}\mathbf{m}|$, where the absolute value $|\cdot|$ is applied to each element, corresponding to the assumption that the absolute values of linear combinations of pixels have an entropy distribution. This filtering allows correlations between pixels to be taken into account.

The choice of the conditional distribution $p(\mathbf{g}|\mathbf{f})$ depends on both the imaging model and the model for the noise. If we assume spatially uncorrelated additive Gaussian statistics for the noise, we find the following for the conditional distribution:

$$p(\mathbf{g}|\mathbf{f}) = C \exp\left[-\frac{1}{2\sigma^2} \|\mathbf{H}\mathbf{f} + \mathbf{b} - \mathbf{g}\|^2\right], \quad (7)$$

where σ^2 is the variance of the noise. Note that we can eliminate the known or estimated background \mathbf{b} by subtracting it from \mathbf{g} . Therefore we omit \mathbf{b} in the equations for Gaussian noise statistics.

In many situations the assumption of Poisson statistics is more appropriate than that of Gaussian statistics. The measured image is described as a photon-counting process:

$$p(\mathbf{g}|\mathbf{f}) = \prod_{i=1}^M \frac{\mu_i^{N_i} \exp(-\mu_i)}{N_i!}, \quad (8)$$

where $\mu_i = \beta[\mathbf{H}\mathbf{f} + \mathbf{b}]_i$ is the mean, $N_i = \beta\mathbf{g}_i$ is the number of photons, and β is the reciprocal of the photon-conversion factor. For fluorescence microscopy the photon-conversion factor is determined by several physical parameters, such as integration time and the quantum efficiency of the detector.¹² The background \mathbf{b} cannot be incorporated into a term that is independent of \mathbf{f} and needs to be taken into account explicitly.

We substitute Eq. (7) or (8) and Eq. (4) or (5) into Eq. (3) and take the negative of the logarithm, yielding log-likelihood functionals that need to be minimized to produce MAP estimates. Leaving out terms that are independent of \mathbf{f} leads to four possible functionals Ψ . They are as follows:

for Gaussian noise statistics and a Gaussian prior:

$$\Psi_{G,G} = \|\mathbf{H}\mathbf{f} - \mathbf{g}\|^2 + \gamma\|\mathbf{C}(\mathbf{f} - \mathbf{m})\|^2; \quad (9)$$

for Poisson noise statistics and a Gaussian prior:

$$\Psi_{P,G} = \sum \mathbf{H}\mathbf{f} - \mathbf{g}^T \ln(\mathbf{H}\mathbf{f} + \mathbf{b}) + \gamma\|\mathbf{C}(\mathbf{f} - \mathbf{m})\|^2, \quad (10)$$

where the sum is over all elements of the vector $\mathbf{H}\mathbf{f}$ and the function $\ln(\cdot)$ is applied to each element of its argument;

for Gaussian noise statistics and an entropy prior:

$$\Psi_{G,E} = \|\mathbf{H}\mathbf{f} - \mathbf{g}\|^2 + \gamma|\mathbf{C}\mathbf{f}|^T \ln\left(\frac{|\mathbf{C}\mathbf{f}|}{e|\mathbf{C}\mathbf{m}|}\right), \quad (11)$$

where divisions of vectors are done element by element;

for Poisson noise statistics and an entropy prior:

$$\Psi_{P,E} = \sum \mathbf{H}\mathbf{f} - \mathbf{g}^T \ln(\mathbf{H}\mathbf{f} + \mathbf{b}) + \gamma|\mathbf{C}\mathbf{f}|^T \ln\left(\frac{|\mathbf{C}\mathbf{f}|}{e|\mathbf{C}\mathbf{m}|}\right). \quad (12)$$

Equations (9)–(12) are of the general form $F + \gamma P$, in which F measures the quality of the fit of \mathbf{f} to the data and the penalty function P measures the roughness of \mathbf{f} . Minimizing this type of functional is known in regularization theory as the stabilization functional approach.¹³

The model \mathbf{m} incorporates prior knowledge of the shape of the object. The regularization matrix \mathbf{C} is used to extract those features of the solution that need to be penalized. The regularization parameter γ determines the trade-off between the accuracy of the fit and the amount

of smoothing. The MAP approach for deriving Eqs. (9)–(12) yields the following choices of γ :

$$\gamma_{G,G} = \sigma^2/\tau^2, \quad (13)$$

$$\gamma_{P,G} = 1/(2\beta\tau^2), \quad (14)$$

$$\gamma_{G,E} = 2\sigma^2\rho, \quad (15)$$

$$\gamma_{P,E} = \rho/\beta. \quad (16)$$

Under certain circumstances we might wish to choose γ differently and therefore retain an unspecified γ in the further development of the algorithms. The practical problems associated with the selection of γ , \mathbf{C} , and \mathbf{m} will be discussed further in Subsection 2.E.

The interpretation of Eqs. (9)–(12) as a regularization approach gives us a practical reason for specifying a Gaussian distribution as the prior distribution function of \mathbf{f} . The quadratic choice of P in Eqs. (9) and (10) is commonly used to produce stable solutions.¹³

C. Discussion of Related Work

The functionals that were given in the previous subsection are the starting point for the development of a number of restoration algorithms. Before proposing new strategies, we give a short summary of related work. The literature in this area is vast, and we restrict ourselves to some of the more relevant references.

Equation (9) is also known as the Tikhonov–Miller functional.^{14,15} The unconstrained minimum of this functional can be obtained analytically. In the presence of a non-negativity constraint, two approaches have been used. Carrington and co-workers applied a transformation of the functional and unconstrained minimization to find a nonnegative solution.^{5,6} The iterative constrained Tikhonov–Miller algorithm combines a conjugate gradient minimization with projection.^{7,8}

The EM algorithm² has been studied extensively to find a solution that maximizes Eq. (8). In the astronomy community this algorithm is known as the Richardson–Lucy method and has been widely used for restoring astronomical images.¹⁶ Holmes applied the algorithm for fluorescence microscopy³ and presented a modification to improve the slow convergence rate.⁴

Snyder *et al.*¹⁷ showed that maximizing the mean value of the log likelihood of Eq. (8) is equivalent to minimizing Csiszár's I divergence. Using a set of axioms of regularity, Csiszár¹⁸ argued that the I divergence is the only appropriate distance measure for nonnegative functions. The I divergence $I(\mathbf{x}, \mathbf{y})$ is equal to $-S(\mathbf{x}, \mathbf{y})$, the negative of the entropy given by Eq. (6). Maximizing Eq. (8) is indeed equivalent to maximizing the entropy of \mathbf{g} relative to $\mathbf{H}\mathbf{f} + \mathbf{b}$.

MAP formulations of the image restoration have been proposed in the literature. Trussell and Hunt formulated a MAP image restoration algorithm for Gaussian prior and conditional distributions.¹⁹ The entropy distribution function has been used as a prior distribution for the object.²⁰ Hunt^{9,10} proposed a number of algorithms for MAP image restoration, one of which solves Eq. (10) with $\gamma = 1$ and $\mathbf{C} = \mathbf{I}$. Another algorithm involves the

use of the Poisson distribution as a prior density for the object. This yields the following functional to be minimized:

$$\Psi = \sum \mathbf{H}\mathbf{f} - \mathbf{g}^T \ln(\mathbf{H}\mathbf{f}) - \mathbf{f}^T \ln \mathbf{m} + \sum \ln(\mathbf{f}!). \quad (17)$$

Before derivatives of this equation can be computed, the factorial in $\ln(\mathbf{f}_i!)$ must be simplified. Stirling's approximation gives

$$\ln(\mathbf{f}_i!) \approx -\mathbf{f}_i + \mathbf{f}_i \ln \mathbf{f}_i + \frac{1}{2} \ln(2\pi\mathbf{f}_i). \quad (18)$$

Hunt argues that, for typical imaging conditions, $\ln(\mathbf{f}_i!) \approx \mathbf{f}_i \ln \mathbf{f}_i$. Note, however, that the alternative approximation $\ln(\mathbf{f}_i!) \approx -\mathbf{f}_i + \mathbf{f}_i \ln \mathbf{f}_i$ yields a functional that is equal to Eq. (12) with $\mathbf{b} = \mathbf{0}$, $\gamma = 1$, and $\mathbf{C} = \mathbf{I}$. Hunt uses algorithms similar to the EM algorithm to find the MAP solutions.

D. Proposed Algorithms

Equations (9)–(12) must be minimized subject to the constraint that \mathbf{f} be nonnegative. We implement this constraint by making the transformation of variables $\mathbf{f}_i = \mathbf{x}_i^2$. Unconstrained minimization with respect to \mathbf{x} and backtransformation yield the nonnegative minimum. Let \mathbf{x}^2 be given by \mathbf{x}_i^2 , \mathbf{X} be a diagonal matrix with $\mathbf{X}_{ii} = \mathbf{x}_i$, and \mathbf{S} be a diagonal matrix with $\mathbf{S}_{ii} = \text{sgn}[(\mathbf{C}\mathbf{x}^2)_i]$. The gradients of the transformed functionals with respect to \mathbf{x} are given by

$$\nabla \Psi_{G,G} = 4\mathbf{X}[\mathbf{H}^T(\mathbf{H}\mathbf{x}^2 - \mathbf{g}) + \gamma\mathbf{C}^T\mathbf{C}(\mathbf{x}^2 - \mathbf{m})], \quad (19)$$

$$\nabla \Psi_{P,G} = 2\mathbf{X} \left[\mathbf{H}^T \left(1 - \frac{\mathbf{g}}{\mathbf{H}\mathbf{x}^2 + \mathbf{b}} \right) + 2\gamma\mathbf{C}^T\mathbf{C}(\mathbf{x}^2 - \mathbf{m}) \right], \quad (20)$$

$$\nabla \Psi_{G,E} = 4\mathbf{X} \left\{ \mathbf{H}^T(\mathbf{H}\mathbf{x}^2 - \mathbf{g}) + \frac{1}{2} \gamma\mathbf{C}^T\mathbf{S} \left[\ln \left(\frac{|\mathbf{C}\mathbf{x}^2|}{e|\mathbf{C}\mathbf{m}|} \right) + 1 \right] \right\}, \quad (21)$$

$$\nabla \Psi_{P,E} = 2\mathbf{X} \left\{ \mathbf{H}^T \left(1 - \frac{\mathbf{g}}{\mathbf{H}\mathbf{x}^2 + \mathbf{b}} \right) + \gamma\mathbf{C}^T\mathbf{S} \left[\ln \left(\frac{|\mathbf{C}\mathbf{x}^2|}{e|\mathbf{C}\mathbf{m}|} \right) + 1 \right] \right\}, \quad (22)$$

where divisions of vectors are done element by element. Setting these derivatives equal to zero and solving for \mathbf{x} yield the square root of the MAP estimate.

Note that all four derivatives are proportional to \mathbf{x} . Thus any element of \mathbf{x} that is equal to zero is not changed by an iterative algorithm based on searches along directions that are linear combinations of these gradients. That is, the spatial extent of the solution is limited to those elements of \mathbf{x} that are nonzero in the initial estimation, providing a means for constraining the spatial support of the solution.

Analytical solutions for the roots of Eqs. (19)–(22) are not available. Appendix A describes how iterative solutions for \mathbf{f} can be obtained by using the conjugate gradient algorithm. We implemented this approach in the C pro-

gramming language. In the rest of this paper, we denote the algorithms that minimize Eqs. (19), (20), (21), and (22) as MAPGG, MAPPG, MAPGE, and MAPPE, respectively. Using Eqs. (19) and (20) with $\gamma = 0$, we formulated ML algorithms for the cases of Gaussian noise and Poisson noise. We refer to them as MLG and MLP, respectively.

E. Choice of the Regularization Parameter

The choice of the regularization parameter γ is critical for a good restoration result, since it determines the trade-off between fitting and smoothing the solution. Ideally, the choice of γ should depend on some objective measure on the restoration result. However, this is difficult without knowledge of the true object. The problem of the choice of the regularization parameter for the case of Gaussian priors and noise statistics in the absence of further constraints has been studied extensively,^{21–23} since the MAP problem is then linear and the solution is known analytically. The choice of γ is far more complicated for nonlinear MAP problems.

Equations (13)–(16) are the choices that follow from the MAP formulation of the restoration problem. The variance σ^2 and the photon-conversion factor $1/\beta$ are physical parameters that, in principle, could be estimated from the image. Thus the regularization parameter should be chosen proportional to the amount of noise in the image determined by σ^2 and β .

The choices of τ and ρ are more difficult. If \mathbf{C} is the identity matrix and \mathbf{f}_i is well described by independent Gaussian-distributed variables, then an obvious choice for τ is the well-known estimator for the standard deviation of multidimensional uncorrelated Gaussian variables, $\|\hat{\mathbf{f}} - \mathbf{m}\|/\sqrt{M}$, where $\hat{\mathbf{f}}$ is some estimation of the object. With this choice of τ , and for the special case $\mathbf{m} = 0$, Eq. (13) becomes equivalent to expressions that are derived with different reasoning by Miller¹⁵ and Katsaggelos *et al.*²⁴ In Subsection 2.B we noted that Eq. (9) can be written as a weighted sum of a fitting term F and a penalty P . We then see that this choice of τ normalizes P such that it has the same scale as that of F . With this in mind we propose to normalize the regularization parameters such that choosing them interactively becomes more convenient. Setting $\tau \propto \|\mathbf{C}(\hat{\mathbf{f}} - \mathbf{m})\|$ in Eqs. (13) and (14) and $\rho \propto -S(|\mathbf{C}\hat{\mathbf{f}}|, |\mathbf{C}\mathbf{m}|)^{-1}$ in Eqs. (15) and (16) leads to penalty terms with the same scale as that of the fitting terms. In summary, we choose the regularization parameters as follows:

$$\gamma_{G,G} = \kappa \frac{M\sigma^2}{\|\mathbf{C}(\hat{\mathbf{f}} - \mathbf{m})\|^2}, \quad (23)$$

$$\gamma_{P,G} = \kappa \frac{M}{2\beta\|\mathbf{C}(\hat{\mathbf{f}} - \mathbf{m})\|^2}, \quad (24)$$

$$\gamma_{G,E} = 2\kappa M\sigma^2 \left[|\mathbf{C}\hat{\mathbf{f}}|^T \ln \left(\frac{|\mathbf{C}\hat{\mathbf{f}}|}{e|\mathbf{C}\mathbf{m}|} \right) + \sum |\mathbf{C}\mathbf{m}| \right]^{-1}, \quad (25)$$

$$\gamma_{P,E} = \kappa \frac{M}{\beta} \left[|\mathbf{C}\hat{\mathbf{f}}|^T \ln \left(\frac{|\mathbf{C}\hat{\mathbf{f}}|}{e|\mathbf{C}\mathbf{m}|} \right) + \sum |\mathbf{C}\mathbf{m}| \right]^{-1}, \quad (26)$$

where κ is chosen interactively. The factor M normalizes the regularization parameter with respect to the size of the image. The remaining problem is the choice of $\hat{\mathbf{f}}$. We found that $\hat{\mathbf{f}} = \mathbf{g}$ is satisfactory. Using experience from restorations of similar images and using visual feedback on the result, one can vary κ to obtain an optimal result. An objective means for selecting the regularization parameter γ is, of course, much more desirable and remains a topic of future research.

The choice of the regularization matrix \mathbf{C} is also problematical, inasmuch as it determines which features of \mathbf{f} are penalized. In image restoration applications one usually resorts to a high-pass filter, e.g., the Laplace operator. Note, however, that care must be taken for \mathbf{C} in Eqs. (11) and (12). A Laplace filter in combination with a flat model \mathbf{m} would result in divisions by zero. A better choice would be a filter that preserves the lowest frequency, the total sum of the intensities. The sum of a Laplace filter and an identity matrix is one possibility. In the absence of prior knowledge, we might simply set $\mathbf{C} = \mathbf{I}$, also simplifying numerical computations.

The model \mathbf{m} can be used to incorporate prior knowledge of the shape of the object. In the absence of such knowledge, we chose $\mathbf{m} = \mathbf{0}$ in the Gaussian priors and $\mathbf{m} = \text{constant}$ (flat model) in the entropy priors, again simplifying numerical computations.

3. SIMULATIONS AND EXAMPLES

A. Simulated Confocal Imaging

To investigate the properties of our algorithms and compare them with those of existing algorithms, we performed simulations, representative of imaging in a laser scanning confocal microscope. An object was generated on a $64 \times 64 \times 32$ sampling lattice, with a sampling interval of 25 nm in the X – Y direction and 75 nm in the Z direction. The image contained two objects: (1) a hollow sphere, centered on the middle of the image, with a diameter of 1125 nm and a thickness of 37.5 nm; and (2) a sphere of diameter 375 nm, centered 200 nm to the left of the middle of the image. The intensity of the large sphere was renormalized to 200 units. The intensity of the small sphere was renormalized to 100 units.

The simulated confocal image was obtained by convolving the object by a calculated PSF²⁵ with the following parameters: excitation wavelength 488 nm, emission wavelength 520 nm, NA 1.4, and refractive index 1.515. The radius of the pinhole, backprojected in the object plane, was 250 nm. A constant background with intensity 10 was added. Figures 1(a) and 1(b) show the object and the blurred object, respectively.

Gaussian or Poisson noise with varying variance or photon-conversion factor was generated in the image. For Gaussian noise we define the signal-to-noise ratio (SNR) as

$$\text{SNR} = \frac{\max_i (\mathbf{Hf})_i - \min_i (\mathbf{Hf})_i}{\sigma}, \quad (27)$$

where \mathbf{Hf} is the blurred object. Such a simple definition is not possible in the case of Poisson noise. However, the Poisson noise is completely characterized by the photon-conversion factor $1/\beta$. In our experiments the mean of

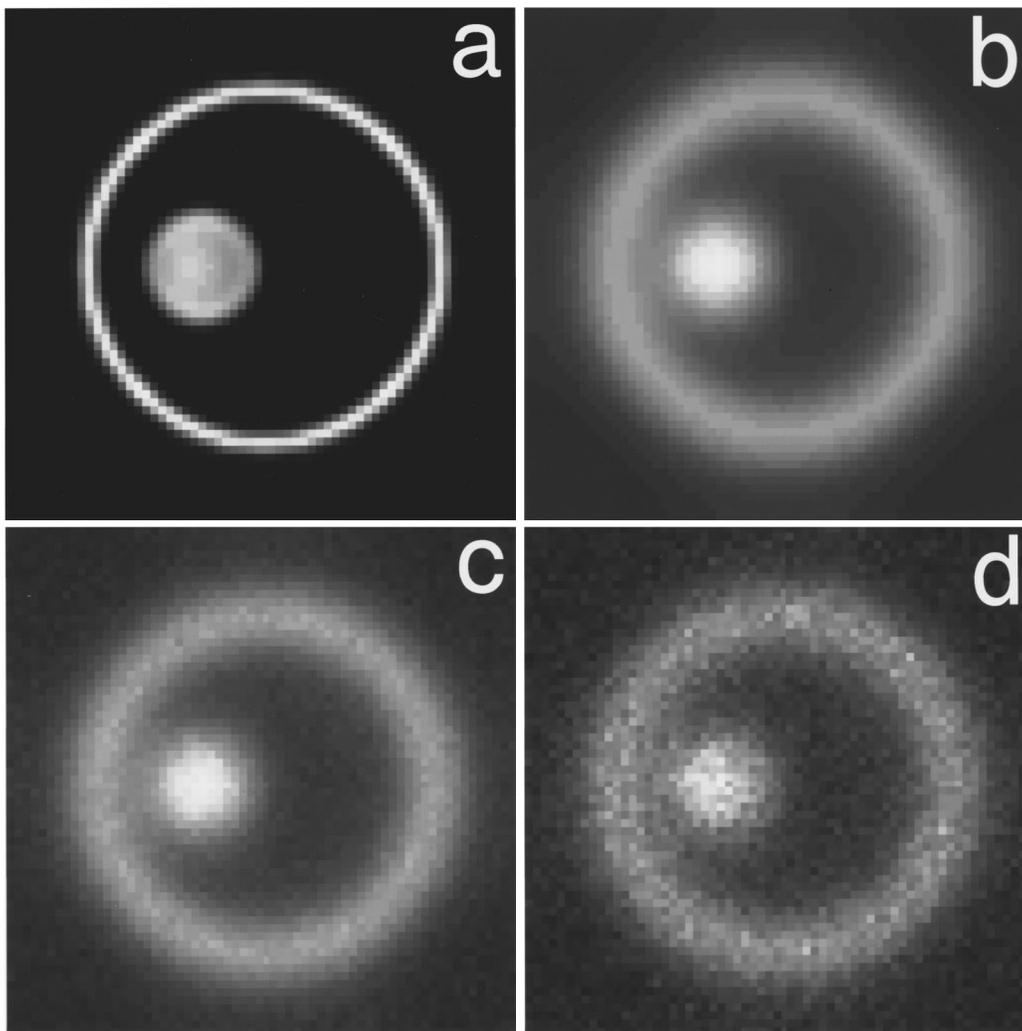


Fig. 1. Objects that are used in the simulations: (a) test object; (b) test object, blurred by simulated confocal PSF; (c) blurred object distorted with Gaussian noise with SNR = 45; (d) blurred object distorted with Poisson noise with $\beta = 10$. A single slice from the center of each stack is shown.

the Poisson process is chosen equal to $\beta \mathbf{H}\mathbf{f}$, where β was varied. As a result, the number of detected photons increases if β increases, yielding higher signals. Thus β acts as a SNR measure.

The criterion for comparing restoration results depends on the application. In Ref. 26 different criteria were used to compare restoration algorithms acquired by confocal microscopy. In this paper we restrict ourselves to the mean-square-error criterion defined by

$$\text{MSE}(\hat{\mathbf{f}}, \mathbf{f}) = \frac{1}{M} \sum_{i=1}^M (\hat{\mathbf{f}}_i - \mathbf{f}_i)^2. \quad (28)$$

In Subsection 2.E we stated the importance of the choice of the regularization parameter and proposed a normalization method to ease the interactive choice of this regularization parameter. To study the quality of the restoration as a function of the regularization parameter for our four MAP restoration algorithms, we computed the mean square error (MSE) for varying values of κ . We

generated a test image with Gaussian noise with SNR = 45 [Fig. 1(c)] and a test image with Poisson noise with $\beta = 10$ [Fig. 1(d)]. The MAPGG and MAPGE algorithms were applied to the image with Gaussian noise, and the MAPPG and MAPPE algorithms were applied to the image with Poisson noise. We used Eqs. (23)–(26) to choose the regularization parameter for given κ . For the estimation $\hat{\mathbf{f}}$ in these equations, the noisy image was used, precorrected by subtracting the constant background. This was also the choice for the initial estimation of the algorithms. We set $\mathbf{m} = \mathbf{0}$ in the MAPGG and MAPPG algorithms, $\mathbf{m} = \mathbf{1}$ in the MAPGE and MAPPE algorithms; and $\mathbf{C} = \mathbf{I}$. This reflects our lack of knowledge about the shape of the object and the correlations between the pixels of the sampled object.

In Fig. 2 the MSE is plotted for the four MAP algorithms as a function of κ . Three important observations can be made from this plot: (1) For the MAPGE and MAPPE algorithms, the regularization parameter needs to be at least 1 order of magnitude larger than that for the MAPGG and MAPPG algorithms. (2) The MAPGE algo-

rithm yields a lower MSE than the MAPGG algorithm, and the MAPPE algorithm yields a lower MSE than the MAPPG algorithm. That is, entropy priors yield lower errors, measured according to the MSE criterion, than

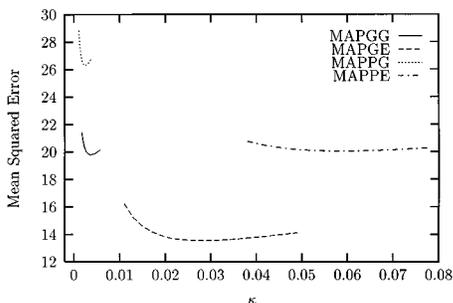


Fig. 2. MSE as a function of κ for a simulated confocal image. Poisson noise with $\beta = 10$ has been added for the MAPPG and MAPPE algorithms. Gaussian noise with SNR = 45 has been added for the MAPGG and MAPGE algorithms. For all four algorithms 500 iterations were used to find the MSE for given κ .

Gaussian priors. (3) Underestimating the regularization parameter leads to larger errors than overestimating it by the same amount. Note that no conclusion should be drawn from the fact that the MAPPG and MAPPE algorithms produce lower errors than the MAPGG and MAPGE algorithms, respectively, inasmuch as these results applied to different images with different types of noise. The values for κ that yielded minimal values in these plots were used in all further simulations.

Figure 3 shows the center slices of the restoration results of the simulations described above, with the use of optimal values of κ . From the simulations that we describe below, we find that the MLG and MLP algorithms do not converge to a useful solution, since they are not regularized. The results depend strongly on the number of iterations. For the MAP algorithms an optimal result for each algorithm could be ensured by choosing the number of iterations (500) high enough for convergence to a fixed point and by choosing an optimal κ from Fig. 2. For this reason we show the results only for the MAP algorithms and do not compare results for the MLG and MLP

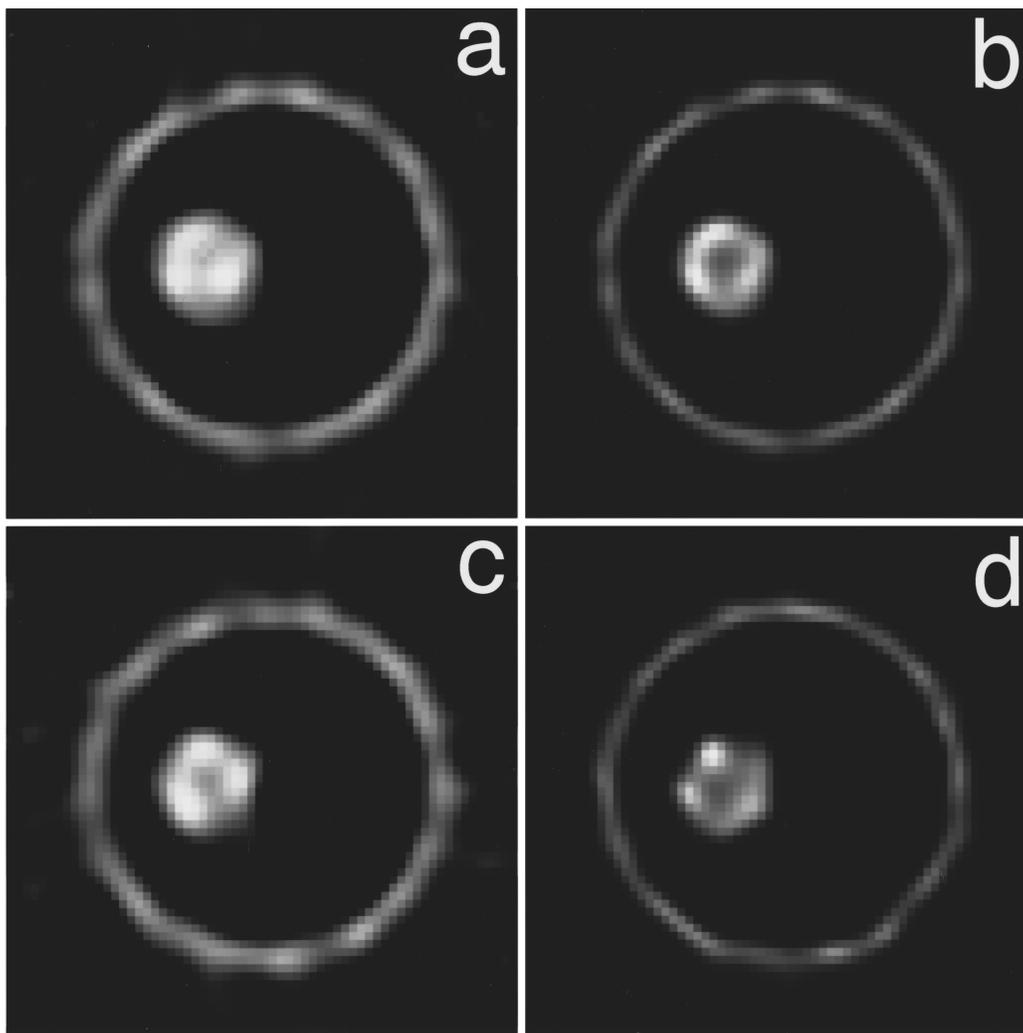


Fig. 3. Results for the MAP algorithms for the simulated images with use of the optimal κ found from Fig. 2: (a) MAPGG result with $\kappa = 3.7^{-3}$; (b) MAPGE result with $\kappa = 2.9^{-2}$; (c) MAPPG result with $\kappa = 2.5^{-3}$; (d) MAPPE result with $\kappa = 5.6^{-2}$. Results (a) and (b) are for the image with Gaussian noise, and results (c) and (d) are for the image with Poisson noise. In all four cases 500 iterations were used. A single slice from the center of each stack is shown.

algorithms visually. When the Gaussian prior [Figs. 3(a) and 3(c)] is used, the hollow sphere is too thick and small artifacts have appeared in the background. When the entropy prior is used, a much sharper result is obtained and no artifacts are visible in the background. However, the smaller sphere is somewhat shrunken compared with the original object.

To analyze the behavior of the algorithms, we computed the MSE as a function of the number of iterations. The same images were used as those in the simulations described above. Figure 4 shows the MSE as a function of the number of iterations for the MAPGG, MAPGE, MLG, and Carrington's algorithms and for the image that contained Gaussian noise. Carrington's algorithm aims to minimize Eq. (9) with $\mathbf{C} = \mathbf{I}$ and $\mathbf{m} = \mathbf{0}$. The regularization parameter was therefore chosen equal to that used for the MAPGG algorithm. For a high number of iterations, the MAPGG, MAPGE, and Carrington's algorithms converge to a fixed point, and the values of the MSE also converge. The MLG algorithm is not regularized and does not converge to a fixed point. As a result, the MSE increases dramatically for a high number of iterations. After convergence the MSE for the MAPGG algorithm is larger than that for the MAPGE algorithm. Note, however, that for a lower number of iterations the values of the MSE for the MAPGG, MAPGE, and MLG algorithms reach approximately the same minimal value. Thus, although the MAPGE algorithm yields lower error at convergence, both the MAPGG and MLG algorithms can attain a similar low value if the number of iterations is selected appropriately. Although Carrington's algorithm aims to do the same as the MAPGG algorithm, it searches the solution space by using a completely different path. As a result, the MSE increases for a low number of iterations before decreasing again to reach the same value as that of the MAPGG algorithm. Thus, in contrast to the MAPGG, MAPGE, and MLG algorithms, Carrington's algorithm cannot be terminated safely before convergence in this example. For different objects and/or different choices of the regularization parameter, this behavior is not always present. We observed cases in which the mean-square-error results were more similar to those for the MAPGG algorithm.

Figure 5 shows the MSE as a function of the number of iterations for the MAPPG, MAPPE, and MLP algorithms and the accelerated EM algorithm of Holmes. The MAPPG, MAPPE, and MLP algorithms exhibit a qualitative behavior similar to that shown in Fig. 4 for the MAPGG, MAPGE, and MLG algorithm. The MSE of the accelerated EM algorithm decreases slowly to the approximate same minimum value as that from the MAPPG, MAPPE, and MLP algorithms. For a high number of iterations, the MSE increases again for the accelerated EM algorithm. Like the MLP algorithm, the accelerated EM algorithm finds the ML solution. However, since the accelerated EM algorithm is much slower, it is easier to select a number of iterations that yields a good restoration result.

To investigate the effect of Gaussian or entropy priors in MAP estimations, we studied the MSE as a function of the amount of noise in the image. Figure 6 presents the MSE as a function of the SNR for images with Gaussian

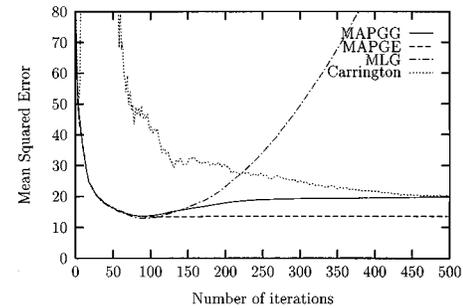


Fig. 4. MSE as a function of the number of iterations for the MAPGG, MAPGE, MLG, and Carrington algorithms. A simulated confocal image was used with Gaussian noise with $\text{SNR} = 45$. The same values for κ were used as those given in Fig. 3 for the MAP algorithms. For Carrington's algorithm the same regularization parameters were used as those for the MAPGG algorithm.

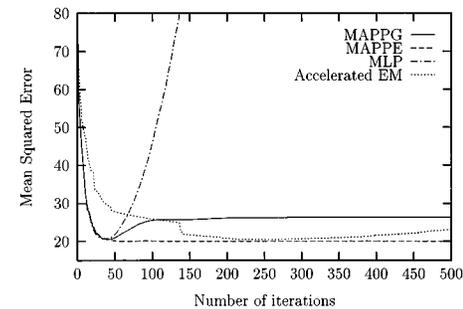


Fig. 5. MSE as a function of the number of iterations for the MAPPG, MAPPE, and MLP algorithms and the accelerated EM algorithm of Holmes. A simulated confocal image was used with Poisson noise with $\beta = 10$. The same values for κ were used as those given in Fig. 3 for the MAP algorithms.

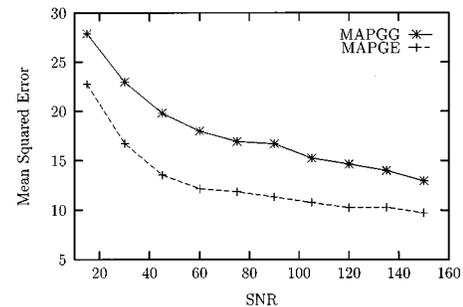


Fig. 6. MSE as a function of the SNR for the MAPGG and MAPGE algorithms. A simulated confocal image was used with Gaussian noise with varying SNR. The values for κ were the same as those given in Fig. 3. The number of iterations was 500.

noise for the MAPGG and MAPGE algorithms. The entropy prior yields a consistently lower error than that from the Gaussian prior. The same behavior can be seen for the case of Poisson noise in Fig. 7, in which the MSE is plotted as a function of the reciprocal of the photon-conversion factor. Note that for increasing SNR or β , the error decreases. For each algorithm, κ was chosen once, with the use of Fig. 2. For different values of SNR or β , the regularization parameter was computed by using Eqs.

(23)–(26) with this value of κ . The results in Figs. 6 and 7 indicate that these choices for the regularization parameter work well.

To determine the computational load of the algorithms, we measured the average time per iteration derived from 500 iterations for each algorithm. Table 1 lists the average time per iteration for the MLG, MAPGG, MAPGE, MLP, MAPPG, MAPPE, accelerated EM, and Carrington's algorithms. The simulated objects were the same as those used in the first two simulations described above. The simulations were done on a Transtec workstation with a 167-MHz UltraSparc processor, running the Solaris 2.5 operating system. Clearly, the use of the entropy prior comes at a computational cost. Use of Gaussian priors in the MAPGG and MAPPG algorithms adds little or no computation time compared with that with use of the ML algorithms MLG and MLP. The MLP, MAPPG, and MAPPE algorithms are slower than the MLG, MAPGG, and MAPGE algorithms, respectively, since the former require one matrix multiplication more than the latter to compute the gradient of the functional. The MAPGG algorithm requires the same amount of computation time as Carrington's algorithm, but as we have seen in the simulations described above, the number of iterations required for Carrington's algorithm might be much larger. The MLP algorithm requires slightly more computation time per iteration than the Holmes algorithm to find the ML estimation. However, we conclude from our simulations that the MLP algorithm converges in a much smaller number of iterations. In summary,

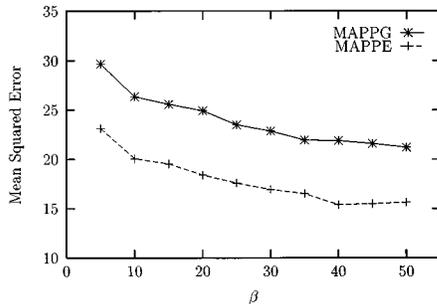


Fig. 7. MSE as a function of β , the reciprocal of the photon-conversion factor of the Poisson noise for the MAPPG and MAPPE algorithms. A simulated confocal image was used with Poisson noise with varying β . The values for κ were the same as those given in Fig. 3. The number of iterations was 500.

Table 1. Average Execution Times per Iteration on a Transtec Workstation with a 167-MHz UltraSparc Processor, Running the Solaris 2.5 Operating System

Algorithm	Average Execution Time (s)
MLG	1.0
MAPGG	1.0
MAPGE	1.7
Carrington	1.0
MLP	1.7
MAPPG	1.8
MAPPE	4.7
Accelerated EM	1.4

our approach to finding MAP estimations for image restoration applications seems to be computationally efficient compared with existing approaches.

B. Application to Confocal Imaging

We applied our algorithms to confocal images of Vero cells, a monkey kidney tumor cell line, which were exposed to cholera toxin and incubated with an antibody specific for the A subunit of the toxin and labeled with the sulfoindocyanine dye Cy5.²⁷ The images were acquired with a Zeiss LSM10 confocal laser scanning microscope, with the following parameters: voxel size $160 \times 160 \times 200$ nm, NA 1.4, excitation wavelength 633 nm, emission wavelength 690 nm, and a backprojected pinhole radius of 338 nm.

We assumed that the noise could be described by a Poisson process and estimated from flat regions in the image that $\beta \approx 10$. We also assumed a constant background and estimated its value at $b \approx 20$. A calculated PSF²⁵ was used. The MAPPG and MAPPE algorithms were selected, since these are optimal for the case of Poisson noise. The MAPPG algorithm was applied with $\kappa = 1$, and the MAPPE algorithm was applied with $\kappa = 5$. These values were chosen relatively high to ensure adequate noise suppression. The value for the MAPPE algorithm was higher than that for the MAPPG algorithm, in accordance with the simulations described above, which indicate that κ should be chosen higher for the entropy algorithms. In both cases 25 iterations were used. With more than 25 iterations the solutions ceased to vary appreciably, indicating that the results were close to the global minimum. Figure 8 shows regions of an optical section from the center of the original stack [part (a)], the MAPPG result [part (b)], and the MAPPE result [part (c)]. The MAPPE result is sharper than the MAPPG result. For 3-D representations and further analysis of the images, see Bastiaens *et al.*²⁷

4. DISCUSSION AND CONCLUSIONS

We have presented algorithms for computing MAP image restorations of images that are blurred and contaminated by Gaussian or Poisson noise. Our approach for minimizing the MAP functionals is computationally very efficient. Since we were motivated by applications in fluorescence microscopy, we compared our algorithms with two others in common use.

We developed two types of algorithm, optimized for Gaussian and Poisson noise statistics, respectively. The Poisson model is the most appropriate model for fluorescence microscopy.²⁸ However, if the number of detected photons is high, the Poisson density function may be approximated by a Gaussian curve. Furthermore, it is important to note that all algorithms incorporate a nonnegativity constraint. This rejects physically impossible negative solutions that may follow from the Gaussian noise model. Indeed, recent work^{26,29} indicates that for higher SNR's, algorithms assuming Gaussian statistics and using a nonnegativity constraint may perform almost as well as algorithms based on Poisson statistics. Considering that the former are much faster than the latter, they may provide the more attractive alternative.

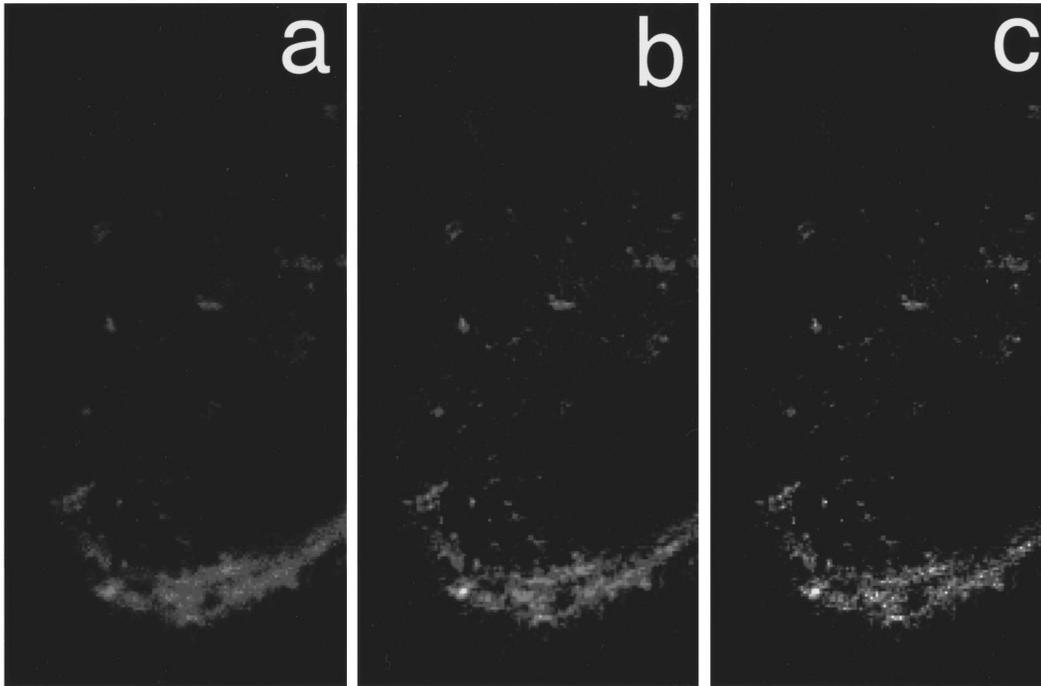


Fig. 8. Confocal image of a Vero cell, treated with cholera toxin and incubated with Cy5-labeled antibody against the A subunit of the toxin, and restoration results. (a) original; (b) MAPPG restoration with $\kappa = 1$; (c) MAPPE restoration with $\kappa = 5$. In both (a) and (b), 25 iterations were used. Part of an optical section from the center of each stack is shown.

A major issue is the choice of the regularization parameter in the MAP algorithms. We propose a normalization method to ease the choice of the regularization parameter if the variance of the noise in the image is known. In practical applications this is usually not the case. It might then be possible to estimate the noise level from the image.³⁰ Even then the choice of κ must be made interactively, based on experience with previous restoration on similar images and using visual feedback on the result. Clearly, further research is necessary to find reliable methods choosing the regularization parameter. Recent work²⁹ indicates that methods developed for the linear Tikhonov filter (which is the unconstrained solution for the MAP problem with Gaussian statistics and prior) might be useful for the MAPPG algorithm. Results from maximum entropy methods³¹ could possibly be adapted for use with our methods that use entropy priors. We did not discuss in depth the choices of the model \mathbf{m} and the regularization matrix \mathbf{C} in this paper. Selection of optimal \mathbf{m} and \mathbf{C} is probably problem dependent and requires further investigation.

We found in our simulations that although algorithms using the entropy prior yield the lowest error according to the criterion of MSE, it is possible to achieve the same low value with algorithms that use the Gaussian prior or the ML algorithms by stopping the iterative algorithms at an early stage. A good method for determining the optimal number of iterations should be developed.

ACKNOWLEDGMENTS

This work was supported by project ERBCHRXCT930177 of the European Community, "Human capability and mobility: fluorescence *in situ* hybridization and image

analysis." We thank P. I. H. Bastiaens for providing the images of Vero cells and G. M. P. van Kempen for many discussions on image restoration. We thank W. A. Carington for providing valuable information about his algorithm.

APPENDIX A: MINIMIZATION ALGORITHMS

The nonlinear conjugate gradient algorithm is used to minimize Eqs. (9)–(12). The algorithm is given at iteration k by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{d}^k, \quad (\text{A1})$$

where

$$\mathbf{d}^k = \beta^k \mathbf{d}^{k-1} - \nabla \Psi(\mathbf{x}^k) \quad (\text{A2})$$

and β^k is given by

$$\beta^k = \frac{\|\nabla \Psi(\mathbf{x}^k)\|^2}{\|\nabla \Psi(\mathbf{x}^{k-1})\|^2}. \quad (\text{A3})$$

The step α^k along the search direction \mathbf{d}^k is found by minimizing the functional $\Psi(\mathbf{x} + \alpha \mathbf{d})$ for all values of α .

For a functional of the form $\Psi = (\mathbf{x}^2)^T \mathbf{A} \mathbf{x}^2 - 2(\mathbf{x}^2)^T \mathbf{q} + c$, this is equivalent to minimizing a fourth-order polynomial:

$$\Psi(\mathbf{x} + \alpha \mathbf{d}) = p\alpha^4 + q\alpha^3 + r\alpha^2 + s\alpha + t, \quad (\text{A4})$$

where

$$p = (\mathbf{d}^2)^T \mathbf{A} \mathbf{d}^2, \quad (\text{A5})$$

$$q = 4(\mathbf{d}^2)^T \mathbf{A} \mathbf{X} \mathbf{d}, \quad (\text{A6})$$

$$r = 4\mathbf{d}^T \mathbf{X} \mathbf{A} \mathbf{X} \mathbf{d} + 2(\mathbf{d}^2)^T (\mathbf{A} \mathbf{x}^2 - \mathbf{q}), \quad (\text{A7})$$

$$s = 4\mathbf{d}^T\mathbf{X}(\mathbf{A}\mathbf{x}^2 - \mathbf{q}), \quad (\text{A8})$$

$$t = (\mathbf{x}^2)^T\mathbf{A}\mathbf{x}^2 - 2(\mathbf{x}^2)^T\mathbf{q} + c. \quad (\text{A9})$$

The extrema of a fourth-order polynomial can be found analytically.

For the case of Gaussian statistics and a Gaussian prior, $\mathbf{A} = \mathbf{H}^T\mathbf{H} + \gamma\mathbf{C}^T\mathbf{C}$, $\mathbf{q} = \mathbf{H}^T\mathbf{g} + \gamma\mathbf{C}^T\mathbf{C}\mathbf{m}$, and $c = \mathbf{g}^T\mathbf{g} + \gamma\mathbf{m}^T\mathbf{C}^T\mathbf{C}\mathbf{m}$.

For the other three functionals, it is not possible to optimize α^k analytically. However, we can derive the first and second derivatives of $\Psi(\mathbf{x} + \alpha\mathbf{d})$ with respect to α . Newton's rule can be applied to find the zero crossings of the derivatives by using the second derivatives. A few iterations of Newton's rule are sufficient. We always use three iterations. The iterations are computationally not very expensive if the matrix multiplications are computed before applying Newton's rule. For the case of Poisson statistics with a Gaussian prior, we find

$$\begin{aligned} \frac{d\Psi_{P,G}}{d\alpha} &= 2(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{H}^T \left[1 - \frac{\mathbf{g}}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right] \\ &+ \gamma(4p\alpha^3 + 3q\alpha^2 + 2r\alpha + s), \end{aligned} \quad (\text{A10})$$

$$\begin{aligned} \frac{d^2\Psi_{P,G}}{d\alpha^2} &= 2(\mathbf{d}^2)^T\mathbf{H}^T \left[1 - \frac{\mathbf{g}}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right] \\ &+ 4\mathbf{g}^T \left[\frac{\mathbf{H}(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right]^2 \\ &+ \gamma(12p\alpha^2 + 6q\alpha + 2r), \end{aligned} \quad (\text{A11})$$

where p , q , r , and s are given by Eqs. (A5)–(A8), where $\mathbf{A} = \mathbf{C}^T\mathbf{C}$ and $\mathbf{q} = \mathbf{C}^T\mathbf{C}\mathbf{m}$.

For the case of Gaussian statistics with an entropy prior, we get

$$\begin{aligned} \frac{d\Psi_{G,E}}{d\alpha} &= 4p\alpha^3 + 3q\alpha^2 + 2r\alpha + s \\ &+ 2\gamma(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \left\{ 1 + \ln \left[\frac{|\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2|}{e|\mathbf{C}\mathbf{m}|} \right] \right\}, \end{aligned} \quad (\text{A12})$$

$$\begin{aligned} \frac{d^2\Psi_{G,E}}{d\alpha^2} &= 12p\alpha^2 + 6q\alpha + 2r + 2\gamma(\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \\ &\times \left\{ 1 + \ln \left[\frac{|\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2|}{e|\mathbf{C}\mathbf{m}|} \right] \right\} \\ &+ 4\gamma(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \frac{\mathbf{C}(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)}{\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2}, \end{aligned} \quad (\text{A13})$$

where the diagonal matrix \mathbf{S} is given by $\mathbf{S}_{ii} = \text{sgn}([\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2]_i)$ and p , q , r , and s are given by Eqs. (A5)–(A8), with the use of $\mathbf{A} = \mathbf{H}^T\mathbf{H}$ and $\mathbf{q} = \mathbf{H}^T\mathbf{g}$.

For the case of Poisson noise statistics and an entropy prior, we find

$$\begin{aligned} \frac{d\Psi_{P,E}}{d\alpha} &= 2(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{H}^T \left[1 - \frac{\mathbf{g}}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right] \\ &+ 2\gamma(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \\ &\times \left\{ 1 + \ln \left[\frac{|\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2|}{e|\mathbf{C}\mathbf{m}|} \right] \right\}, \end{aligned} \quad (\text{A14})$$

$$\begin{aligned} \frac{d^2\Psi_{P,E}}{d\alpha^2} &= 2(\mathbf{d}^2)^T\mathbf{H}^T \left[1 - \frac{\mathbf{g}}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right] \\ &+ 4\mathbf{g}^T \left[\frac{\mathbf{H}(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)}{\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})^2 + \mathbf{b}} \right]^2 + 2\gamma(\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \\ &\times \left\{ 1 + \ln \left[\frac{|\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2|}{e|\mathbf{C}\mathbf{m}|} \right] \right\} \\ &+ 4\gamma(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)^T\mathbf{C}^T\mathbf{S} \frac{\mathbf{C}(\mathbf{X}\mathbf{d} + \alpha\mathbf{d}^2)}{\mathbf{C}(\mathbf{x} + \alpha\mathbf{d})^2}. \end{aligned} \quad (\text{A15})$$

Address all reprint requests to Thomas M. Jovin at the location on the title page; tel, 49-551-2011382; fax, 49-551-2011467. E-mail addresses, peterv@mpc186.mpibpc.gwdg.de and tjovin@mpc186.mpibpc.gwdg.de.

REFERENCES

1. D. A. Agard, Y. Hiraoka, P. Shaw, and J. W. Sedat, "Fluorescence microscopy in three dimensions," in *Methods in Cell Biology*, D. L. Taylor and Y. Wang, eds. (Academic, San Diego, Calif., 1989), Vol. 3, pp. 353–377.
2. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B* **39**, 1–38 (1977).
3. T. J. Holmes, "Maximum-likelihood image restoration adapted for noncoherent optical imaging," *J. Opt. Soc. Am. A* **5**, 666–673 (1988).
4. T. J. Holmes and Y. Liu, "Acceleration of maximum-likelihood image restoration for fluorescence microscopy and other noncoherent imagery," *J. Opt. Soc. Am. A* **8**, 893–907 (1991).
5. W. Carrington, "Image restoration in 3D microscopy with limited data," in *Bioimaging and Two-Dimensional Spectroscopy*, L. C. Smith, ed., Proc. SPIE **1205**, 72–83 (1990).
6. W. A. Carrington, R. M. Lynch, E. D. W. Moore, G. Isenberg, K. E. Fogarty, and F. S. Fay, "Superresolution three-dimensional images of fluorescence in cells with minimal light exposure," *Science* **268**, 1483–1487 (1995).
7. R. L. Lagendijk, "Iterative identification and restoration of images," Ph.D. dissertation (Delft Technical University, Delft, The Netherlands, 1990).
8. H. T. M. van der Voort and K. C. Strasters, "Restoration of confocal images for quantitative image analysis," *J. Microsc.* **178**, 165–181 (1995).
9. B. R. Hunt, "Prospects for image restoration," *Int. J. Mod. Phys. C* **5**, 151–178 (1994).
10. B. R. Hunt, "Super-resolution of images: algorithms, principles, performance," *Int. J. Imaging Syst. Technol.* **6**, 297–304 (1995).
11. J. Skilling, "Classic maximum entropy," in *Maximum Entropy and Bayesian Methods*, J. Skilling, ed. (Kluwer Academic, Dordrecht, The Netherlands, 1989), pp. 45–52.
12. T. M. Jovin and D. J. Arndt-Jovin, "Luminescence digital imaging microscopy," *Annu. Rev. Biophys. Chem.* **18**, 271–308 (1989).
13. M. Bertero and P. Boccacci, "Regularization methods in im-

- age restoration: an application to HST images," *Int. J. Imaging Syst. Technol.* **6**, 376–386 (1995).
14. A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems* (Wiley, New York, 1977).
 15. K. Miller, "Least-squares method for ill-posed problems with prescribed bound," *SIAM (Soc. Ind. Appl. Math.) J. Math. Anal.* **1**, 52–74 (1970).
 16. H.-M. Adorf, R. N. Hook, and L. B. Lucy, "HST image restoration developments at the ST-ETF," *Int. J. Imaging Syst. Technol.* **6**, 339–349 (1995).
 17. D. L. Snyder, J. Schulz, and J. A. O'Sullivan, "Deblurring subject to nonnegativity constraints," *IEEE Trans. Signal Process.* **40**, 1143–1150 (1992).
 18. I. Csizsár, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *Ann. Stat.* **19**, 2033–2066 (1991).
 19. H. J. Trussell and B. R. Hunt, "Improved methods of maximum a posteriori restoration," *IEEE Trans. Comput.* **C-27**, 57–62 (1979).
 20. K. M. Hanson, "Bayesian and related methods in image reconstruction from incomplete data," in *Image Recovery: Theory and Application*, H. Stark, ed. (Academic, San Diego, Calif., 1987), pp. 79–125.
 21. N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Process.* **1**, 322–336 (1992).
 22. S. J. Reeves and R. M. Mersereau, "Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration," *Opt. Eng.* **29**, 446–454 (1990).
 23. A. M. Thompson, J. C. Brown, J. W. Kay, and D. M. Titterton, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 326–339 (1991).
 24. A. K. Katsaggelos, J. Biemond, R. W. Schafer, and R. M. Mersereau, "A regularized iterative image restoration algorithm," *IEEE Trans. Acoust. Speech Signal Process.* **39**, 913–929 (1991).
 25. H. T. M. van der Voort and G. J. Brakenhoff, "3-D image formation in a high-aperture fluorescence confocal microscope: a numerical analysis," *J. Microsc.* **158**, 43–54 (1990).
 26. G. M. P. van Kempen, H. T. M. van der Voort, J. G. J. Bauman, and K. C. Strasters, "Comparing maximum likelihood estimation and constrained Tikhonov–Miller restoration," *IEEE Eng. Med. Biol. Mag.* **15**, 76–83 (1996).
 27. P. I. H. Bastiaens, I. V. Majoul, P. J. Verveer, H.-D. Söling, and T. M. Jovin, "Imaging the intracellular trafficking and state of the AB₅ quaternary structure of cholera toxin," *EMBO J.* **15**, 4246–4253 (1996).
 28. C. J. R. Sheppard, X. Gan, M. Gu, and M. Roy, "Signal-to-noise in confocal microscopes," in *Handbook of Biological Confocal Microscopy*, 2nd ed., J. B. Pawley, ed. (Plenum, New York, 1995), pp. 363–371.
 29. G. M. P. van Kempen, L. J. van Vliet, P. J. Verveer, and H. T. M. van der Voort, "A quantitative comparison of image restoration methods for confocal microscopy," *J. Microsc.* **185**, 354–365 (1997).
 30. S. I. Olsen, "Estimation of noise in images: an evaluation," *Comput. Vis. Graph. Image Process.* **55**, 319–323 (1993).
 31. S. F. Gull, "Developments in maximum entropy data analysis," in *Maximum Entropy and Bayesian Methods*, J. Skilling, ed. (Kluwer Academic, Dordrecht, The Netherlands, 1989), pp. 53–71.