

Image Restoration for Confocal Microscopy: Improving the Limits of Deconvolution, with Application to the Visualization of the Mammalian Hearing Organ

Jacques Boutet de Monvel, Sophie Le Calvez, and Mats Ulfendahl

Institute for Hearing and Communication Research, Karolinska Institute, Karolinska Hospital, SE-171 76 Stockholm, Sweden

ABSTRACT Deconvolution algorithms have proven very effective in conventional (wide-field) fluorescence microscopy. Their application to confocal microscopy is hampered, in biological experiments, by the presence of important levels of noise in the images and by the lack of a precise knowledge of the point spread function (PSF) of the system. We investigate the application of wavelet-based processing tools to deal with these problems, in particular wavelet denoising methods, which turn out to be very effective in application to three-dimensional confocal images. When used in combination with more classical deconvolution algorithms, these methods provide a robust and efficient restoration scheme allowing one to deal with difficult imaging conditions. To make our approach applicable in practical situations, we measured the PSF of a Biorad-MRC1024 confocal microscope under a large set of imaging conditions, including in situ acquisitions. As a specific biological application, we present several examples of restorations of three-dimensional confocal images acquired inside an intact preparation of the hearing organ. We also provide a quantitative assessment of the gain in quality achieved by wavelet-aided restorations over classical deconvolution schemes, based on a set of numerical experiments that we performed with test images.

INTRODUCTION

Thanks to its optical sectioning properties, laser scanning confocal microscopy is an invaluable tool for the visualization of complex biological structures in three dimensions (for a general reference, see Pawley, 1995). To take the example that will be considered in this article, the mammalian hearing organ, the organ of Corti, is enclosed inside the cochlea, a tube-like structure divided into different compartments, which in turn is embedded in the temporal bone. Confocal microscopy has proven extremely useful to bypass the difficulty of access inside the temporal bone and to study this system under conditions close to normal living conditions (Flock et al., 1998, 1999; Ulfendahl et al., 2000). The limited access inside a nearly intact inner ear, however, constrains us to use suboptimal imaging conditions. When focusing deep inside the organ of Corti, much of the light is lost due to scattering by surrounding structures, and various distortions may occur due to variations of the optical index inside the specimen. A subsequent processing of the images is required to partly remedy these problems.

Different image restoration methods have been designed over the years. A popular method is based on the application of deconvolution algorithms, allowing one to compensate for the blurring caused by out-of-focus contributions to the measured intensity at each pixel of the image. Such a blurring effect occurs in any imaging process to some

extent, but it is particularly important to consider in biological microscopy, where the interesting details are often of a size comparable to the system's resolution. For this reason, deconvolution has become increasingly popular in fluorescence microscopy, with the advent of laser scanning systems and of personal computers capable of processing huge three-dimensional (3D) images.

Although deconvolution algorithms have proven very effective in conventional (wide-field) fluorescence microscopy, their application to biological confocal microscopy is hampered by two main difficulties. One is that images acquired in a confocal setting are quantum limited: Each pixel receives only a limited and random number of photons per unit time, with most of the out-of-focus light being rejected by the small confocal aperture. As a consequence, noise problems are usually important in confocal microscopy and must be treated with particular care (Pawley, 1995). A second, even more important difficulty is due to the lack of a precise control on the imaging conditions in biological experiments. When focusing light deep inside a complex specimen, one can usually hope for only an approximate knowledge of the point spread function (PSF) of the system, which is the crucial information needed for deconvolution. These problems, together with a common belief that confocal images, being already deblurred, cannot be further improved, seem to have limited the use of deconvolution in confocal microscopy. This method is more commonly applied in conventional microscopy, where intensity levels are significantly higher and accurate measurements of the PSF are easier to obtain and more stable from one sample to another.

This article is concerned with the processing methods that can be used to improve on the limits imposed by the two problems mentioned above. We shall focus on methods

Received for publication 10 November 1999 and in final form 26 December 2000.

Address reprint requests to Dr. Jacques Boutet de Monvel, ENT Research Laboratory and Institute for Hearing and Communication Research, M9-ENT Karolinska Hospital, SE-171 76 Stockholm, Sweden. Tel.: 46-8-5177-3210; Fax: 46-8-348546; E-mail: j.boutet.de.monvel@ihk.ki.se.

© 2001 by the Biophysical Society

0006-3495/01/05/2455/16 \$2.00

based on wavelet analysis, in particular wavelet denoising algorithms, which appear to be most relevant to 3D fluorescence microscopy. We put an emphasis on their use in laser fluorescence confocal microscopy, but we stress that the processing tools that we describe are not restricted to this particular type of microscopy. They are in particular relevant to two-photon and to wide-field imaging systems as well. These methods are of real biological interest, because they allow one to obtain results of a quality that would not be possible without increasing the specimen exposure to radiation.

The paper is organized as follows. In the next section, we explain the experimental and theoretical basis for deconvolution applied to confocal imaging systems. We describe different ways to determine the PSF of a confocal microscope and report on a thorough experimental investigation that we have performed with a Biorad-MRC1024 confocal microscope. We then review the Bayesian approach to deconvolution in the form suitable for confocal microscopy. Such methods are effective when the PSF of the system is perfectly known but appear to be seriously limited when this is not the case. In the third section we describe and discuss the application of wavelet denoising (Donoho and Johnstone, 1994), which can be used as an efficient regularization of deconvolution and lead to significant improvement of the restorations in practical situations. The important question of a quantitative assessment of these methods is discussed in the fourth section, where we report on a set of experiments performed on test images, clearly demonstrating the improvement achieved by wavelet-aided restorations over more classic methods. We finally conclude by pointing out some of the issues remaining to be solved, together with some directions of research that seem to us promising for the future.

DECONVOLUTION FOR CONFOCAL MICROSCOPY

Determination of the point spread function of a confocal microscope

Before one can use deconvolution algorithms, it is crucial to get a reasonably precise knowledge of the imaging system's PSF. One must make a distinction between the PSF measured under imaging conditions close to optimal with respect to the system's design, and the PSF measured *in situ*, i.e., under the imaging conditions of biological experiments. The former PSF is a reproducible characteristic of the microscope, which can be determined in principle from a precise knowledge of the system's optical setup. The latter PSF varies from sample to sample and should be measured on a case-by-case basis. We shall discuss mainly the optimal PSF, as it is much easier to obtain with current systems and it can be considered a reasonable working approximation.

By definition, the PSF of the system is the image that it produces of a point-like source placed at the origin of sample space. Under an ideal confocal setting, this image can be written as the product of two wide-field PSFs, which can in turn be computed for an aberration-free "aplanatic" lens using diffraction theory (Bertero et al., 1990, and references therein). This theoretical model is suitable for describing the bulk PSF of a real confocal microscope, but it does not allow one to capture details important for the sake of deconvolution, such as asymmetries in the tails of the PSF. As these details depend on many parameters, some of which are not easily accessible experimentally, it is important to determine the PSF of a real confocal microscope also by direct measurement.

We used a common experimental approach to measuring PSFs, which consists of imaging fluorescent microspheres of sub-resolution size. With a laser excitation wavelength around 500 nm, beads of 50-nm diameter would provide a safe approximation of point-like sources. However, the images obtained with such small beads have a very poor signal-to-noise ratio (SNR) unless the system is specially optimized. We obtained a better compromise with the Biorad-MRC1024 confocal microscope, by using micro-beads of ~100-nm diameter.

We performed several test experiments to determine the PSF for different sizes of the confocal aperture and under different imaging conditions. Our main results are reproduced in Figs. 1–4. Except for Fig. 4, the beads were immersed in an agarose gel preparation of moderate concentration. Details on our experimental procedure and on the image acquisition settings we used are given at the end of this section. (Further details may be obtained upon request to the authors.)

Fig. 1 shows maximum projections of the PSF and its intensity profile along the optical axis for a theoretical model assuming an ideal confocal setting (with no aberration) and for measurements performed in the 488/585LP excitation/emission filter mode of the system (see the Experimental Methods below for an explanation). Each measurement represents an average over 17–18 different bead images extracted from the raw acquisition, for a given value of the confocal pinhole radius.

We define the full-half-maximum-widths (FHMWs) of the PSF in the standard way, as the intervals $\delta x_{1/2}$, $\delta y_{1/2}$, and $\delta z_{1/2}$, (along the transverse focal axes Ox , Oy , and the optical axis Oz , respectively), where the PSF intensity is more than half of its maximum value. Fig. 2 shows a plot of $\delta x_{1/2}$, $\delta y_{1/2}$, and $\delta z_{1/2}$, measured as a function of the depth of focus below the surface of the gel, keeping the same acquisition settings at different depths. No significant variation of these FHMWs is observed up to ~300 μm below the surface of the gel, and the shape of the PSF was not varying significantly either. At a depth of ~300 μm , $\delta z_{1/2}$ displays a significantly smaller value, but this is an artifact caused by a movement of the gel during the acquisition. Due to this

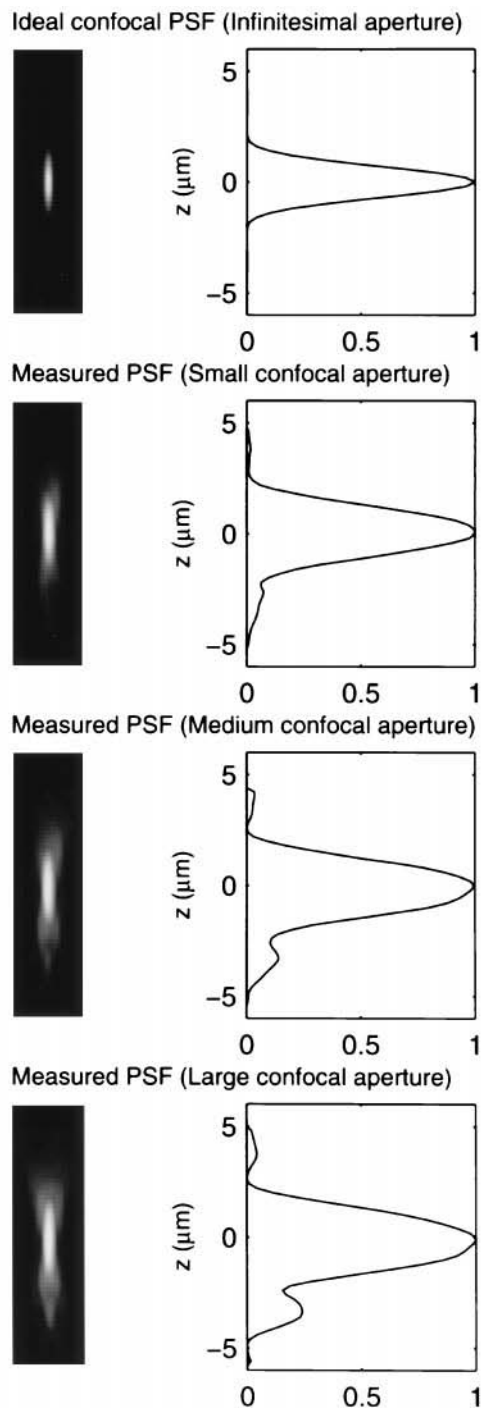


FIGURE 1 Ideal model confocal PSF together with measurements performed on a Biorad-MRC1024 confocal microscope for different openings of the confocal aperture. The measured PSFs represent averages over 17–18 beads, imaged with the 488/585LP excitation/emission filter mode of the system. Each image shows a maximum projection of the PSF parallel to the optical axis, together with the corresponding (normalized) intensity profile along Oz .

movement, the PSF appeared to be abnormally tilted along the Ox direction, which explains why $\delta x_{1/2}$ appears larger, whereas $\delta y_{1/2}$ is not significantly changed.

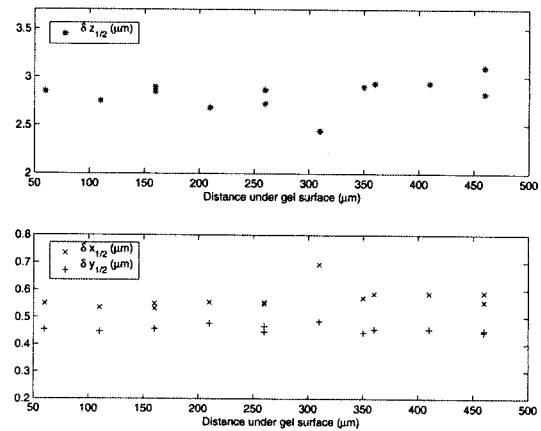


FIGURE 2 Measurements of the FHMWs $\delta x_{1/2}$, $\delta y_{1/2}$, and $\delta z_{1/2}$ (defined in the text) obtained from confocal images of the beads in agarose gel, taken in the 488/605DF32 filter mode under identical acquisition settings, but for different levels of focus under the gel's surface. The anomalous result around 300 μm depth is due to the fact that the gel started to move with respect to the objective lens.

The most important properties of an imaging system as regards to deconvolution are linearity and translation invariance. Linearity (the requirement that a superposition of fluorophore distributions $f_1(x) + f_2(x)$ should produce a superposition of images $I_1(x) + I_2(x)$, $I_i(x)$ being the image of $f_i(x)$, $i = 1, 2$) is not easy to check experimentally because there is no obvious way to superpose two different samples without disturbing them. Although there are known effects limiting the validity of linearity in fluorescence confocal microscopy (mostly absorption in thick specimens, together with saturation and photodegradation of the fluorophore (Pawley, 1995)), it is generally assumed that this property holds to a good accuracy under normal working conditions. Most laser scanning systems used in biological applications operate in a partial or full beam-scanning mode. In the Biorad system the specimen is physically translated along the direction of the optical axis, while the focus point is scanned throughout the focal plane with an arrangement of movable mirrors that slightly deflect the laser beam from the optical axis. Small spatial variations of the PSF are inevitably introduced in this way. However, in most of our experiments we could assume translation invariance to be verified within measurement errors. Significant variations of the PSF occurred only at very high beam deflections, which are rarely explored in practice. Together with the results of Fig. 2, this observation implies that aberrations induced by the objective were very small in our experiments, which is not surprising in view of the high quality of the lens used.

Fig. 3 shows the behavior of the FHMWs of the PSF as the pinhole radius is increased. From this figure and Fig. 1 it is clear that the experimental PSFs are neither symmetric with respect to the focal plane, nor cylindrically symmetric

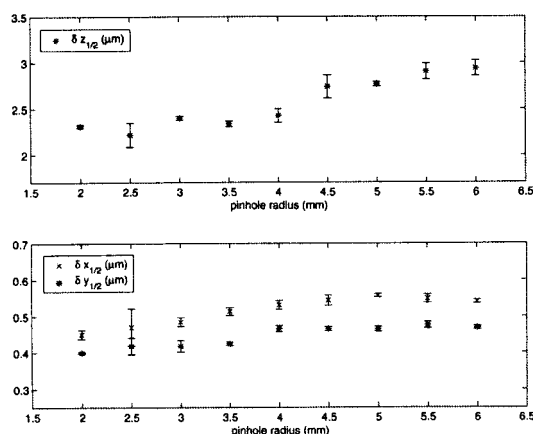


FIGURE 3 Dependence of the FWHMs $\delta x_{1/2}$, $\delta y_{1/2}$, and $\delta z_{1/2}$ upon the size of the confocal aperture (data shown for the 488/605DF32 filter mode). The bars represent standard errors from different (between two and five) confocal images of the beads, taken under the same acquisition settings but at different locations within the gel. Note the slight astigmatism in the PSF $\delta x_{1/2}$ being significantly larger than $\delta y_{1/2}$.

around the optical axis, as $\delta x_{1/2}$ appears systematically larger than $\delta y_{1/2}$. We observed this slight, but clear astigmatism in all our experiments, regardless of excitation wavelength and emission filter used in the acquisition. This effect is most certainly system dependent, and probably reflects non-exact confocality of the excitation spot with respect to the pinhole aperture, due to a slight misalignment of the scanning mirrors of the system. Another possibility would be the presence of slight aberrations or a misalignment of the objective lens. This possibility has to be ruled out, however, because the objective lens was removed and replaced several times during our experiments, and once in place, it could be freely rotated around the optical axis. It is thus clear that the position of the lens was not the same in all our acquisitions; nevertheless, we always observed the same astigmatism.

Finally, Fig. 4 reproduces a limited set of acquisitions obtained by imaging beads attached just above the Reissner's membrane, inside an in vitro preparation of the hearing organ (Ulfendahl et al., 1989). Each PSF in this figure is from a different bead, and the acquisition was performed with a large pinhole aperture (6 mm) to produce an image with a reasonable SNR. One sees in this example that the shape of the PSF is significantly modified by the presence of complex surrounding structures. In particular, the PSF appears tilted along the Ox axis. This was not an artifact caused by a movement of the beads, as was checked by repeating the acquisition several times at the same location. The measured FWHMs of this PSF also appear significantly increased compared with the corresponding optimal PSF. Some of the main features of Figs. 1 and 3 remain; in particular, the same astigmatism effect was observed, with $\delta x_{1/2}$ being larger than $\delta y_{1/2}$ with the same order of discrep-

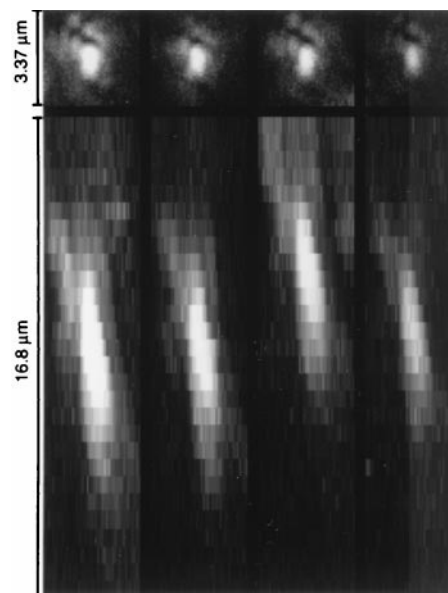


FIGURE 4 Measurements of a confocal PSF in situ, obtained by imaging beads attached to the Reissners membrane inside the inner ear. Each image corresponds to a different bead, without averaging, extracted from an acquisition performed in the 488/585LP filter mode. (Upper part) Maximum projection of the PSF along the optical axis; (Lower part) Maximum projection along the Oy axis. Note the significant distortions in the shape of the PSF, whereas translation invariance is fairly well respected.

ancy. Note that the translation invariance of the PSF is still respected to a fairly good approximation in this example.

EXPERIMENTAL METHODS

Image acquisition and PSF measurements

The Biorad-MRC1024 confocal microscope utilizes a 15-mW krypton/argon laser providing excitation lines at 488 nm and 568 nm, which are transmitted selectively to the specimen through narrow-band (~ 10 -nm) excitation filters. (The system also allows one to exploit the 647-nm line of the laser, but we did not use this line.) For collecting the fluorescence light, we used different spectral emission filters provided by Biorad: a long-pass filter transmitting all wavelengths above 585 nm (referred to as the 585LP filter), a band-pass filter transmitting within the interval $605 \text{ nm} \pm 16 \text{ nm}$ (referred to as the 605DF32 filter), and occasionally a band-pass filter adjusted to the interval $522 \text{ nm} \pm 16 \text{ nm}$ (referred to as the 522DF32 filter). The confocal aperture of the system consists of a circular pinhole of variable radius (ranging between 0.7 and 8 mm), placed in front of a detection photomultiplier (PMT). The size and shape of the PSF of a confocal microscope is primarily determined, for a given objective lens, by the spectrums of excitation and emission wavelengths and by the pinhole radius. Parameters such as the excitation laser intensity and the gain of the PMTs are important to optimize the SNR and dynamic range of the images, but under normal working conditions they do not affect the shape of the PSF significantly.

In experiments, the value of the pinhole radius is usually set by hand (together with the other parameters) to optimize the contrast of the images visually. The PSF of the system must therefore be measured for each excitation/emission filter mode and each value of the pinhole radius used in practice. To keep this feasible we restricted our acquisitions to a representative set of values ranging from 2 to 6 mm by increments of 0.5

mm. When investigating the effect of some other parameter than the pinhole radius, we used a medium radius of 4 mm, as representative of the most common imaging situation for our studies.

Our PSF measurements were performed with Molecular Probes (Eugene, OR) T-8873 TransFluoSpheres, which are microspheres of diameter 0.1 μm , coated with a dye whose excitation peak matches the 488-nm spectral line of the argon laser and whose fluorescence emission is maximal around 605 nm. The beads were imaged with a Zeiss Achromplan 40 \times NA0.75 water-immersion lens. A common experimental procedure for PSF acquisition consists of imaging the beads on a glass slide. Because this method introduces a significant background in the measurements due to reflections by the glass slide, we used another approach where the beads are immersed in an agarose gel of moderate concentration. This allows one to avoid much of the background light and usually leads to better measurements, without greatly complicating the procedure of extraction of the PSF from the images. A slight mismatch exists in this setting between the optical index of the gel and the design index of the objective lens, which might cause aberration effects (Gibson and Lanni, 1991; Hell et al., 1993). These effects should, however, not be significant with a water-immersion lens and if the gel is of reasonably low concentration. For the type of dyes and samples studied in our experiments, we found that the PSF did not depend much on the excitation wavelength or the type of emission filter used. The main parameter determining the size and shape of the PSF was the value of the pinhole radius.

PSF extraction and averaging

Due to the quantum-limited nature of the light collected from a single bead, it is necessary to average the PSF over several different beads to get an accurate measurement. The averaging procedure we used involved four steps: 1) extracting a stack of single-bead images from the raw acquisition, 2) selecting out the bead images representing outliers, which should not be taken into account in the averaging (typical outliers are aggregates of several beads glued together, or couples of different, but close beads that were not separated by the extraction procedure), 3) aligning the selected beads together, and 4) computing the averaged PSF. All these steps may be performed in a large part automatically by using standard image-processing methods. The selection phase is the more involved, however, and it had to be carried out partly by visual inspection. Some care is needed in the alignment procedure, because the center of a given bead has typically non-integer pixel coordinates. (This problem is usually significant only along the optical axis where the resolution is coarsest.) Standard cubic-spline interpolation was used to match the centers of the different beads at half of the pixel resolution, the center of an interpolated bead being determined as the pixel of maximum spatial correlation with some fixed reference PSF. Step 4 is not entirely trivial, because after being selected and aligned properly, the different bead images still have different SNRs and should be weighted accordingly when taking the average. We used the technique described in the Appendix to estimate the SNR for each bead image. Essentially the same results were obtained by simply weighting the bead images by their root-mean-square values, in good agreement with the Poisson nature of the images' statistics. Depending on the beads' density inside the gel, the number of bead images contained in a typical acquisition (of size $\sim 60 \mu\text{m} \times 60 \mu\text{m} \times 12 \mu\text{m}$) ranged between a few dozens to several hundreds. Due to the selection step, however, the number of beads contributing to the final averaged PSF ranged typically between 10 and 30.

Imaging settings in Figs. 1–4

The results shown in this section correspond to 8-bit images acquired in the normal scanning mode of the Biorad system (corresponding to a pixel dwell time a little less than 4 μs), using 10% of the total excitation light intensity of the 15-mW krypton/argon laser, and a sixfold Kalman averaging. Except for Fig. 4, all our acquisitions had format $512 \times 512 \times 31$

with a pixel size of $0.1247 \mu\text{m} \times 0.1247 \mu\text{m} \times 0.4 \mu\text{m}$, corresponding to a $63.8\text{-}\mu\text{m} \times 63.8\text{-}\mu\text{m} \times 12.4\text{-}\mu\text{m}$ region scanned inside the agarose gel. The pixel size of the images shown in Fig. 1 has been divided by two, however, due to the interpolation involved in averaging the PSF. Fig. 4 was extracted from a $512 \times 512 \times 28$ image of pixel size $0.075 \mu\text{m} \times 0.075 \mu\text{m} \times 0.6 \mu\text{m}$, corresponding to a $38.4\text{-}\mu\text{m} \times 38.4\text{-}\mu\text{m} \times 16.8\text{-}\mu\text{m}$ region scanned inside the hearing organ preparation. The measurements of Fig. 1 correspond to identical imaging settings, except for the size of the pinhole radius, which are 2.5 mm, 4 mm, and 6 mm for the small, the medium, and the large aperture, respectively. We used the 488/585LP filter mode, with a PMT gain of 1100 V (this gain is the voltage applied across the dynode chain of the PMT, which can be set, on the Biorad system, between 0 and 1500 V). For the PSF measurements of Fig. 2, we used the 488/605DF32 filter mode, together with a pinhole opening of 4 mm, and a PMT gain of 1050 V. For the measurements shown in Fig. 3 we used the same settings as for Fig. 2, except for the pinhole radius varying from 2 to 6 mm by steps of 0.5 mm. Finally, the images of Fig. 4 were taken in the 488/585LP mode with a large pinhole aperture (7 mm) and a high PMT gain of 1450 V.

Bayesian approach to deconvolution

Commonly used deconvolution algorithms in 3D microscopy include iterative methods, such as the Jansson-van Cittert algorithm (Jansson et al., 1970); algorithms based on Fourier filtering, such as the Wiener filter method and classical least-squares regularization methods based on the minimization of a Tikhonov functional (Tikhonov and Arsenin, 1977; Miller, 1970); and algorithms based on Bayesian inference, the most popular being the Richardson-Lucy algorithm (Richardson, 1972; Lucy, 1974).

Bayesian deconvolution

Both from a theoretical and practical point of view, the Bayesian approach appears the most appropriate for confocal microscopy; it can be based on a physically well founded model of the imaging process and leads to algorithms that are both efficient and easy to implement. We shall briefly review the main ideas underlying this method. First of all, the image restoration problem considered here can be cast as a convolution $I(x) = f * p(x)$, where $f(x)$ represents the local concentration of fluorophore inside the specimen, $p(x)$ is the PSF of the system, and $I(x)$ is the light intensity collected in the image. (In the following, we shall not make an explicit distinction between a volume element $(x, \Delta x)$ of the sample and the corresponding volume element in image space. This amounts to adopting for image space a coordinate system that is conjugate to the one used for sample space, as is usual in microscopy.) This form of the imaging transformation assumes that the system is linear and translation invariant. The PSF is band-limited, so that inversion of the imaging transformation is an ill-posed problem: direct Fourier inversion leads to huge amplification of the low-level, high-frequency components induced by noise in the image. As already pointed out, noise is inevitable in confocal microscopy, and some regularization is needed to avoid deconvolving the noise rather than the signal.

Bayesian restoration makes use of a stochastic model of the imaging process, which specifies the probability $P(I|f)$ to obtain an image $I(x)$ given that the sample fluorophore distribution is $f(x)$. The conditional probability $P(f|I)$ of f given I , also called the posterior probability distribution, is then defined with Bayes' formula $P(f|I) = P(I|f)P(f)/P(I)$. In this formula $P(I)$ serves as a normalization factor (the image I is fixed), and $P(f)$ is the unconditional, or prior, probability distribution of the object f . A certain freedom is usually left on the choice of $P(f)$, but it is an important ingredient of the approach, as it introduces a stochastic constraint on the reconstructed object, which will provide regularization. Once $P(I|f)$ and $P(f)$ are specified, one is interested in deducing some useful information about the object f from the corresponding posterior probability. The max-

imum a posteriori (MAP) estimate $f_{\text{MAP}}(x)$ is obtained by maximization of $P(f|I)$ over all possible values of f , for a given image I . Another natural estimate (more in the spirit of the statistical physics of image restoration (Geman and Geman, 1984; Pryce and Bruce, 1995)) is the posterior mean $E(f(x)|I)$, which is the point-wise average of $f(x)$ over the posterior distribution $P(f|I)$ (formally defined as the integral $E(f(x)|I) = \int f(x)P(f|I)df$ as a function of x). For the purposes of image restoration, the posterior mean has been recognized to be more natural than the MAP estimate, and it often leads to significantly better results (Marroquin et al., 1987; Pryce and Bruce, 1995; Nishimori and Wong, 1999). We shall, however, restrict ourselves to the MAP estimate in the following. The main reason for this is that $f_{\text{MAP}}(x)$ can be computed efficiently by using iterative optimization algorithms, whereas computing $E(f(x)|I)$ involves sampling from the posterior distribution and leads to computationally much more demanding algorithms. Moreover, for the stochastic models considered in this paper, it may be argued that the typical differences between the MAP estimate and the posterior mean are not significant for all practical purposes.

The Richardson-Lucy algorithm

In a laser fluorescence system, the light emission is well approximated by a Poisson point process, where a given volume element Δx of the specimen, centered at x , emits light with an intensity proportional to $f(x)\Delta x$. The image is then also a Poisson point process, the photon count $I(x)$ inside $(x, \Delta x)$ being on average proportional to $f * p(x)\Delta x$. In other words, the likelihood in this model takes the form:

$$P(I|f) = \prod_x \frac{(f * p(x)\Delta x)^{I(x)}}{I(x)!} \exp(-f * p(x)\Delta x). \quad (1)$$

A frequent choice for the prior is the uniform distribution $P(f) = \text{Const}$, assuming no constraint on f . The corresponding MAP estimate, also called the maximum likelihood estimate (MLE), is found by maximizing the second member of Eq. 1 with respect to f . This can be done iteratively by using the Richardson-Lucy (RL) algorithm (Richardson, 1972) (a particular case of the expectation-maximization algorithm (Dempster et al., 1977; Shepp and Vardi, 1982; Holmes, 1988)), which takes the form:

$$f^{k+1}(x) = f^k(x) \left(\left[\frac{I}{f^k * p} \right] * \check{p} \right)(x). \quad (2)$$

In the above expression we have set $\check{p}(x) = p(-x)$, and we assume that normalized units are taken where $\Delta x = 1$ and $\sum_x p(x) = 1$. The expression $f * p(x)$ can then be identified with the discrete convolution $f * p(x) = \sum_y f(y)p(x - y)$, taken as a valid approximation when the pixels are small enough. In addition to its nice physical interpretation, the RL algorithm has a number of robustness properties explaining its popularity in many applications other than confocal microscopy. (It is, for example, widely used in astronomy.) It is, however, a nonregularizing algorithm, so that taking the limit $k \rightarrow \infty$ in Eq. 2 would not lead to any meaningful result. It is desirable in practice to work with a deconvolution algorithm that actually converges to something meaningful. This can be done by choosing the prior distribution $P(f)$ so as to impose a smoothing constraint on the restored image, as is described in the next paragraph.

Introducing a temperature

In the present case it turns out to be useful to impose a maximum-entropy constraint on f . This amounts to taking $P(f)$ in the form $P(f) = \text{Const} \exp(TS(f))$, where the entropy functional $S(f)$ is defined by $S(f) = -\sum_x f(x)(\ln f(x) - 1)$, and T is a positive parameter interpreted as an effective temperature of the image. Besides being empirically relevant (Gull and Daniell, 1978), this choice of $P(f)$ is a natural one when the images are constrained to be positive (Csiszár, 1991). Maximization of the corre-

sponding posterior distribution $P(f|I)$ can be performed iteratively with a straightforward modification of the RL algorithm:

$$f^{k+1}(x) = f^k(x) \left(\left[\frac{I}{f^k * p} \right] * \check{p} \right)(x) - T f^k(x) \ln(f^k(x)). \quad (3)$$

The rate at which these new iterations converge to a solution, and the solution itself, depend strongly on the value of T . For T not too small, a high degree of convergence is attained in a rather short time. At low temperatures, convergence gets much slower, and a large number of iterations may be required before significant differences appear with the $T = 0$ solution. Note that at high temperatures the algorithm becomes unstable, and some relaxation of Eq. 3 is then required to enforce convergence. In interesting cases, however, the relevant temperatures lie well within the frozen region where no stability problem occurs. Pushing Eq. 3 to convergence produces a MAP solution $f_T(x)$, in which the value of T controls the level of restoration. At very low temperatures, Eq. 3 reduces to the RL algorithm, leading to a ground state, which is a nonregularized solution to $f = I * p(x)$. As T is raised the restoration becomes smoother and smoother, and at very high temperatures the solution is essentially constant. ($f_T(x)$ tends to 1 as $T \rightarrow \infty$ for the normalization chosen here.) It is intuitive that a given reasonable measure of the quality of the restoration will be optimized for a certain temperature T_{opt} depending on the problem at hand. Depending on the quality of the data (SNR of the image, accuracy of the PSF used, fineness of the pixel sampling, etc.), more or less regularization is needed. The better the quality, the lower the optimal temperature should be. It is clearly important to dispose of statistical tests of the restoration to perform this choice in a systematic way. We shall return to this point in the fourth section.

WAVELET DENOISING

When we apply an iterative deconvolution algorithm to real 3D images, artifacts eventually develop for several reasons: at some point we begin to deconvolve the noise; we begin to see that we are using a slightly wrong PSF; distortions associated with sampling problems appear, etc. However these artifacts appear in different ways at different scales within the image. In particular, the fine details of an image are more sensitive to noise, mismatches in the PSF, under-sampling, etc., than the larger features. It should therefore be advantageous to treat the different scales separately in the restoration process. Wavelet analysis allows one in some sense to do this, by providing efficient tools for analyzing and processing a signal in a position- and scale-dependent way (for a recent general reference, see Mallat, 1999). Effective algorithms based on the discrete wavelet transform (DWT) have been designed in recent years for estimating a signal from blurred and noisy data (Donoho and Johnstone, 1994; Donoho, 1995; Starck and Bijaoui, 1994). Although the field of application of wavelet-based processing methods is enlarging rapidly, to date their use in image restoration has mostly concerned 1D and 2D situations, e.g., in the analysis of spectra (Fligge and Solanski, 1997), in the restoration of astronomical images (Starck and Bijaoui, 1994), or in application to medical images (Strömberg, 1997). It is quite clear, however, that these methods are of great potential also to 3D microscopy and more generally as

basic tools in the analysis of complex biological structures containing details at many different scales. Here we investigate the use of wavelet denoising algorithms in application to deconvolution for confocal microscopy.

The discrete wavelet transform and wavelet denoising

A basic scheme for multiscale analysis consists in repeated applications of two conjugate scaling transformations $I \rightarrow L.I$ and $I \rightarrow H.I$, where L is a low-pass filter producing a coarse-grained version of an image $I(x)$, and H is high-pass and is formally defined so that $H.I(x) = I(x) - L.I(x)$. Repeated applications of L and H allow one to work at coarser and coarser resolutions, while keeping track of the information when passing from one scale of resolution to the coarser scales. The DWT allows one to implement the above general scheme in a nonredundant way, meaning that the information extracted at different scales and different positions represent independent components of the image (in the same sense that, in a Fourier transform, independent components are extracted from a signal at different frequencies).

Wavelet denoising (Donoho and Johnstone, 1994) exploits the properties of the DWT to recover a signal that has been degraded by the presence of noise. This method consists of applying a DWT to an image, performing some thresholding in the wavelet domain to retain only significant wavelet coefficients, and applying the inverse DWT, thereby producing a denoised image. Noise is present at all scales of the DWT decomposition with essentially the same magnitude in absolute value. In relative value, however, noise is mostly concentrated in the finest scales where it can be easily thresholded out. (More precisely, when the resolution scale is increased by a factor of l , the SNR is increased by a factor of $l^{d/2}$ for a d -dimensional signal.) Significant data, including sharp structures in the image, typically involve sets of high-level wavelet coefficients that are mostly unaffected by the thresholding. In effect, the method produces an adaptive smoothing of the image in which noise is strongly suppressed while sharp features are preserved. Moreover, this is done in a computationally very efficient way.

An important body of results initiated by Donoho and Johnstone (1994) and described in Mallat (1999) shows that in many cases the thresholding procedure described above greatly outperforms more classical denoising methods based on linear filtering. This will be so particularly if one can find a wavelet decomposition that provides a well-adapted, or sparse, representation of the class of signals under study, meaning that a given signal can be approximated accurately with a small number of the building blocks entering in the decomposition. It is still unknown how to design representations that are optimally adapted to the complex geometry of typical images in two and higher dimensions. We may thus expect that the performance of wavelet denoising on

such images is not yet the best of what is theoretically possible. However, in practice this method often provides the state of the art, both in effectiveness and in computational efficacy. In our experiments we used the SURE wavelet shrinkage procedure of Donoho et al. (1995), which we describe in the Appendix. (All the codes used in this article have been written in the Matlab language (The MathWorks, Natick, MA) and are available from the first author.) Fig. 10 in the next section shows the gain in SNR achieved by this method in a set of numerical experiments performed on the test image of Fig. 9.

Application of wavelet denoising to deconvolution

The image restoration scheme that we investigate in this article is based on applying wavelet denoising to the images before a more classical deconvolution algorithm. This approach has several motivations. General multiscale approaches for solving inverse linear problems have been proposed (Donoho, 1995; see also Nowak and Kolaczyk, 2000, which considers inverse problems in the Poisson statistics case). The application of these methods to deconvolution is best understood for 1D signals, wavelets being particularly effective in that case (Mallat, 1999). Significant progress has been made recently on the case of 2D images (Candès and Donoho, 2000); however, this case is still in development, and the situation in 3D and higher dimensions remains to our knowledge largely unexplored. More importantly for our purposes, to date no systematic theory has been proposed to deal with inverse problems for which the inverse kernel is known only approximately, the situation routinely encountered in 3D microscopy. This theoretical difficulty has an important practical consequence: working with an approximate PSF makes it unavoidable to use iterative deconvolution algorithms. For this reason, it is for the moment difficult to propose a wavelet deconvolution scheme that is computationally acceptable for processing huge 3D images. Wavelet denoising, on the other hand, is better understood theoretically and leads to more practical algorithms that achieve very reliable and robust improvements. In view of this state of affairs, it seems to us reasonable to separate the denoising part from the deconvolution part of the restoration. This approach, although not definitive, has the advantage of being immediately applicable and of providing significant results. In effect, wavelet denoising acts as a very efficient regularization scheme for deconvolution. This method does not allow one to achieve a higher resolution in the sense that the frequency bandwidth of the final restored image is not increased. (Achieving this would require using a finer pixel sampling and a more accurate PSF.) Rather, it suppresses many of the artifacts that would be present at a given step of deconvolution without denoising.

Figs. 5–8 show examples of applications of the above restoration scheme to confocal images acquired inside an *in vitro* preparation of the mouse inner ear (Le Calvez and Ulfendahl, 2000). These images were acquired with the Biorad-MRC1024 system, using the same objective lens (Zeiss 40 \times NA0.75 water immersion) as in the PSF experiments described in this paper. Their pixel format, acquisition settings, and biological content are indicated in the legends. To each image, the WD procedure was applied, followed by a MAP deconvolution using the algorithm defined by Eq. 3. The PSFs used for the different cases were determined from a library of pre-measured PSFs obtained from bead acquisitions as described above. (In each case, we used a PSF acquired for the same excitation/emission filter mode and the same pinhole radius as for the image.) Note that some of the images (e.g., the image of Fig. 5) are presumably slightly undersampled along the optical axis. The possible aliasing effects due to undersampling are usually avoided by filtering the analog image above the pixel sampling frequency during the digitalization (Mallat, 1999). This is done at the level of the detector and amounts to imposing that the array of sensors defining the physical pixels does not leave holes in space. Performing deconvolutions of slightly undersampled, but properly filtered images is perfectly sensible; however, this comes at the price of a loss of resolution in the restoration. Despite this, and the fact that our measured PSFs can at best be considered as estimates of the *in situ* PSFs, our restorations lead to a significant increase of detail. The wavelet denoising regularization proves very effective in each case, as the corresponding deconvolutions performed with maximum-entropy regularization alone (not shown) lead to visually much poorer restorations. This will be confirmed on a more quantitative basis in the next section.

Quantitative assessment of image restoration: experiments

An important problem in image restoration is to assess quantitatively the degree of restoration achieved with a deconvolution algorithm. One would like to be able to measure the gain in resolution and reliability in the restored image, as compared with the original acquisition, and to decide what is the best restoration feasible with a given algorithm. The main obstacle to doing this is that the degraded (blurred and noisy) image is the only information available, together with an approximate knowledge of the imaging system's PSF. There is no obvious way of measuring the quality of a proposed restoration unless we have already a strong *a priori* knowledge of the solution, which we rarely have. To date the only way around this problem has been to perform experiments on images blurred with a known realistic model of the imaging system. The study of such artificial cases often leads to insights that turn out to be

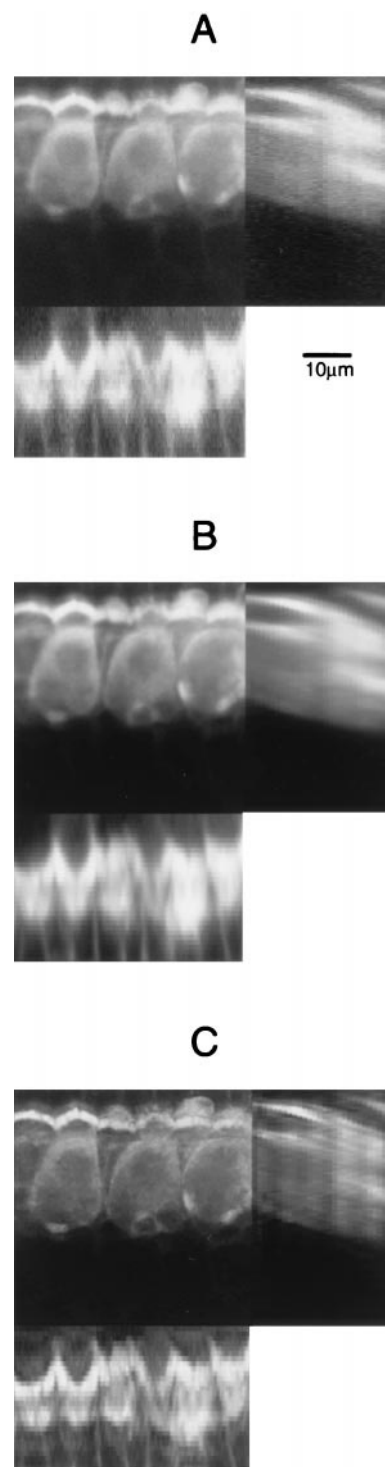


FIGURE 5 Example of wavelet denoising followed by MAP estimation. (A) Sections of a confocal image ($512 \times 512 \times 30$, pixel size $0.0935 \mu\text{m} \times 0.0935 \mu\text{m} \times 1 \mu\text{m}$) showing auditory sensory inner hair cells labeled with the potentiometric styryl dye RH795 (Molecular Probes), which is used to stain cellular membranes. A strong fluorescence is emitted at the level of the sensory hair bundles. In the cellular body, the nucleus and basal vesicles appear darker than the whole cell. Acquisition performed in the 488/585LP excitation/emission filter mode, with a pinhole opening set to 5 mm. (B) The same sections after applying WD to the image in A. (C) Result of a MAP estimation applied on image B ($T = 0.005$, 80 iterations).

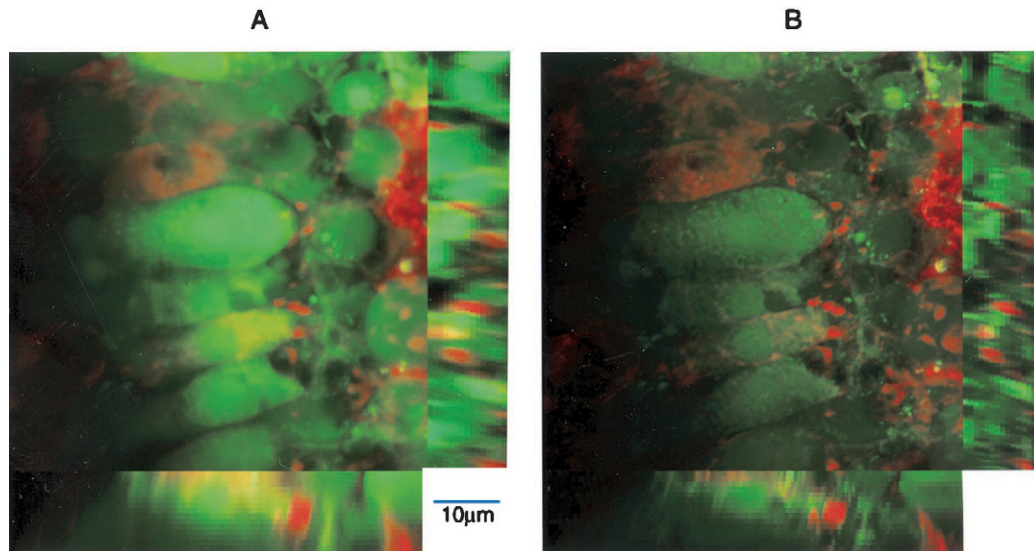


FIGURE 6 (A) Sections of a two-channel confocal image ($512 \times 512 \times 16$, pixel size $0.0935 \mu\text{m} \times 0.0935 \mu\text{m} \times 0.8 \mu\text{m}$) showing auditory inner hair cells at a very apical location in the organ of Corti. The cells are labeled with the RH795 dye (in red) and with calcein (Molecular Probes), which stains uniformly the cytoplasm of living cells (in green). Acquisition performed with filter modes 488/585LP (red channel) and 488/522DF32 (green channel). The pinhole radius was set to 4 mm for both channels. (B) Result obtained by applying WD followed by a MAP estimation with $T = 0.005$ (100 iterations).

very useful in practice to decide what are the optimal parameters (optimal temperature, number of iterations, etc.) under given circumstances.

We investigated these questions using a set of test images such as the one shown in Fig. 9. Fig. 9A represents sections through a 3D sample of pixel format $X \times Y \times Z = 64 \times 64 \times 16$, taken as the undegraded image $f(x)$. The imaging process of a confocal microscope is simulated by writing an

image in the form $I_n(x) = I_0(x) + n(x)$, where $I_0(x) = f * p(x)$ is obtained by blurring the test sample with a known PSF $p(x)$ (see Fig. 11) and $n(x)$ is a noisy component simulating Poisson statistics. In other words, for each pixel x , $I_n(x)$ is a Poisson random variable of mean value $I_0(x)$. Fig. 9, B and C, shows sections of $I_0(x)$ and $I_n(x)$, respectively. Fig. 9D shows the result of applying wavelet denoising to $I_n(x)$ and should be compared with Fig. 9B.

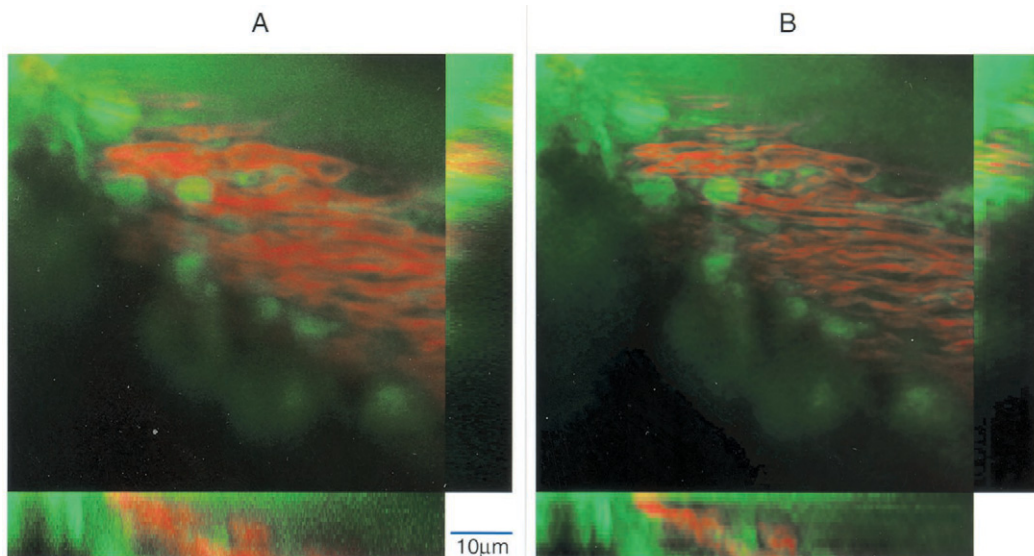


FIGURE 7 (A) Sections of a two-channel confocal image ($512 \times 512 \times 16$, pixel size $0.1247 \mu\text{m} \times 0.1247 \mu\text{m} \times 0.6 \mu\text{m}$) showing thick radial nerve fibers at high magnification, labeled with RH795 (in red) and with calcein (in green). Acquisition performed with the filter modes 488/605DF32 (red channel), and 488/522DF32 (green channel). The pinhole radius was set to 5 mm for the red channel and to 4 mm for the green channel. (B) Result obtained by applying WD followed by a MAP estimation with $T = 0.005$ (150 iterations).

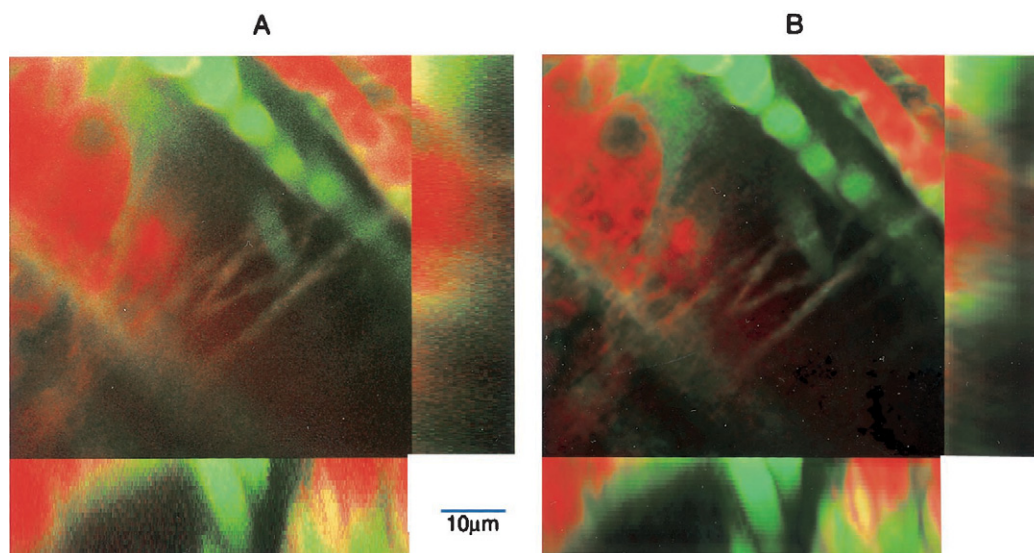


FIGURE 8 (A) Sections of a two-channel confocal image ($512 \times 512 \times 32$, pixel size $0.1247 \mu\text{m} \times 0.1247 \mu\text{m} \times 0.7 \mu\text{m}$) showing upper crossing fibers in the tunnel of Corti. The view presents auditory inner hair cells (left-upper corner in the XY image) emitting a strong red light (corresponding to the RH795 dye), whereas pillar cells appear mainly in green (corresponding to calcein). A small part of the outer hair cell row (right-upper part in the XY image) can be seen in the red labeling. Cellular components like the nucleus and vesicles can be observed in the inner hair cells. Nerve fibers are seen crossing the tunnel of Corti, apparently passing between pillar cells. Acquisition performed with filter modes 488/605DF32 (red channel) and 488/522DF32 (green channel). The pinhole radius was set to 4 mm for both channels. (B) Result obtained by applying WD followed by a MAP estimation at $T = 0.005$ (60 iterations).

Effectiveness of wavelet denoising

Fig. 10 shows the results of a set of denoising experiments performed on the degraded test image (cf. Fig. 9 C) at various levels of noise. The wavelet denoising procedure we have applied is described in the Appendix. For each level of noise, the gain in SNR of the denoising was measured as the ratio $(S/N)_{\text{WD}}/(S/N)$, where S/N and $(S/N)_{\text{WD}}$ denote the SNRs of the image before and after denoising, respectively. Fig. 10 A shows a line-plot of the gain against S/N for a representative set of S/N values. It is clear from Fig. 10 A that the effectiveness of wavelet denoising decreases as S/N increases, in the sense that the gain decreases (from a value of ~ 9.6 at $S/N = 1$ (0 dB) to a value of ~ 2.9 at $S/N \approx 29.5$ (30 dB)). In other words, lowering the level of noise makes it more difficult to distinguish from the signal, which is not surprising. The behavior of the quality factor $Q_2 = 1 - (S/N)/(S/N)_{\text{WD}}$ is also of interest. (Up to a normalization factor, the quantity $(1 - Q_2)^2$ corresponds to the mean-square risk used in the statistical literature.) As shown in Fig. 10 B, Q_2 has a near-linear behavior at small S/N . Linear extrapolation leads to $Q_2 \approx 0.9$ at $S/N = 0$, which corresponds to a maximum gain of ~ 10 . We point out, however, that this extrapolation breaks down at very small S/N and must be considered valid in the present case only for S/N larger than ~ 0.5 (see the remarks at the end of the Appendix).

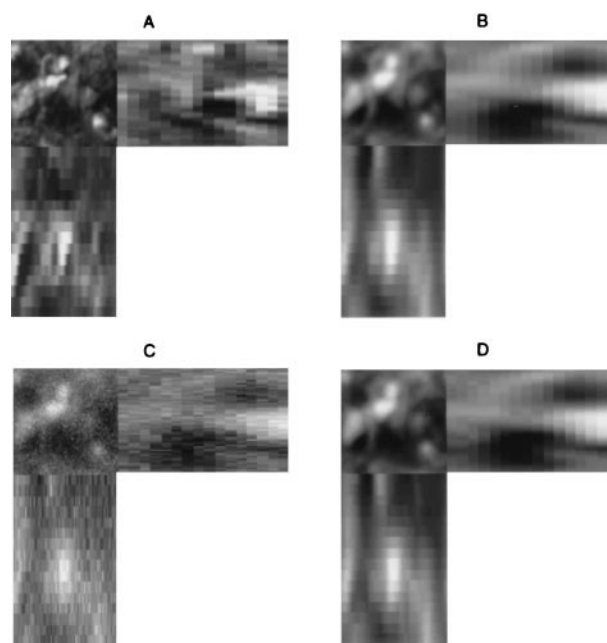


FIGURE 9 The test image used in our numerical experiments. (A) Section through a sample $64 \times 64 \times 16$ image, taken as the original distribution $f(x)$; (B) Result of blurring the image A with the PSF $p(x)$ of Fig. 11; (C) Result of adding a Poisson noisy component to the blurred image B ($S/N = 6.2$); (D) Result of applying wavelet denoising to image C. The gain in SNR is by a factor of 6 ($(S/N)_{\text{WD}} = 37.3$).

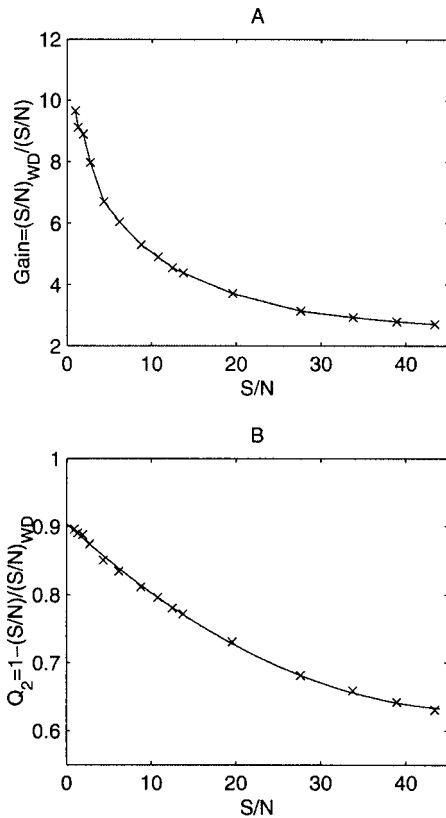


FIGURE 10 Wavelet denoising experiments on the test image of Fig. 9. (A) Line-plot of the gain $(S/N)_{WD}/(S/N)$ against (S/N) , for a set of S/N values in the range $0.97 \leq S/N \leq 43.8$ (corresponding to the decibel interval (0.25 dB, 33 dB)). Note that (S/N) and $(S/N)_{WD}$ are measured in linear units as defined in the Appendix, not in decibels. (B) Behavior of $Q_2 = 1 - (S/N)/(S/N)_{WD}$ as a function of (S/N) . The data points are plotted together with a quadratic fit, extrapolating to $Q_2 \approx 0.90$ at $(S/N) = 0$.

Experiments with the MAP estimate

Figs. 12 and 13 and Table 1 reproduce the results of experiments combining wavelet denoising (WD) and Bayesian deconvolution. Figs. 12 and 13 correspond to two different sets of restorations, applied again to the test image of Fig. 9, at different levels of noise, namely, for $S/N = 4.3, 6.2, 10.7$, and 13.8 . For the first set (Fig. 12), the correct PSF $p(x)$ was used, whereas for the second set (Fig. 13) we used a slightly different function $p_{app}(x)$ (see Fig. 11). For each value of S/N we applied the MAP deconvolution algorithm defined by Eq. 3 to A) the nonprocessed image $I_n(x)$ (see Fig. 9 C) and B) the image obtained by applying WD to $I_n(x)$ (see Fig. 9 D). We shall refer to the restored images obtained in cases A and B as the MAP estimate and the WD+MAP estimate, respectively. In all cases, the MAP or WD+MAP estimate was obtained for a representative set of temperatures around the optimum T_{opt} . Convergence of the deconvolution algorithm was checked by measuring the ratio $\delta_k \Sigma_x (f_{k+1}(x) - f_k(x))^2 / \Sigma_x f_k(x)^2$ at each iteration k . The condition $\delta_k < 10^{-4}$ was used as a stopping criterion. We assessed the

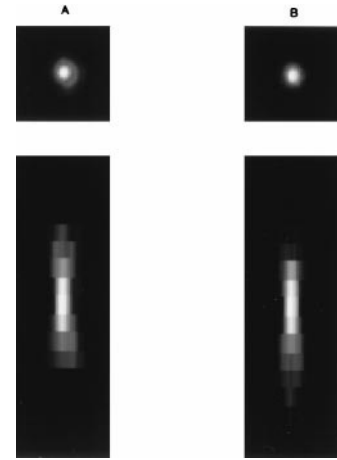


FIGURE 11 Focal and axial projections of the PSFs utilized in the test deconvolutions. (A) The function $p(x)$ used in blurring the test images (referred to as the correct PSF); (B) A slightly different PSF $p_{app}(x)$ (referred to as the approximate PSF $p_{app}(x)$). Both $p(x)$ and $p_{app}(x)$ correspond to measurements (using slightly different optical settings) performed on the Biorad-MRC1024 confocal system.

performance of the restoration by measuring the following root-mean-square quality ratio as a function of T :

$$Q_2 = 1 - \left(\frac{\sum_x (f(x) - f_T(x))^2}{\sum_x (f(x) - I_n(x))^2} \right)^{1/2}, \quad (4)$$

where $f_T(x)$ denotes the MAP or WD+MAP estimate corresponding to temperature T . Parts A and B of Figs. 12 and 13 are line-plots of Q_2 as a function of T for the bare MAP estimate and the WD+MAP estimate, respectively. (Figs. 12, C and D, and 13, C and D, are for the moment out of our

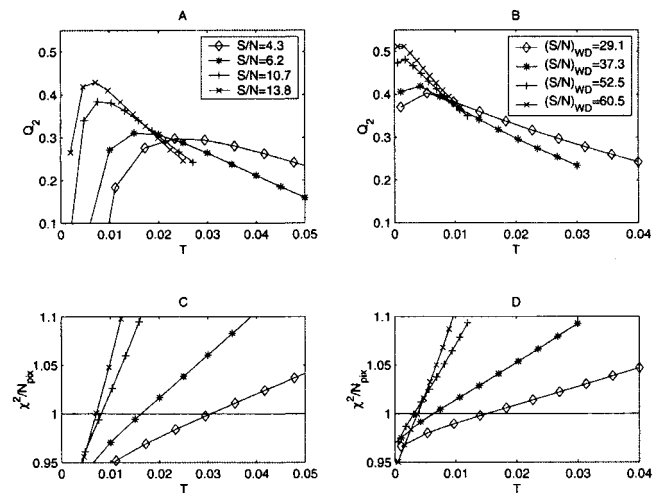


FIGURE 12 Results of MAP and WD+MAP restorations performed on the test image of Fig. 9 (exact PSF used at $S/N = 4.3, 6.2, 10.7$, and 13.8). (Upper part) Plots of Q_2 as a function of T for the MAP estimate (A) and the WD+MAP estimate (B). (Lower part) Plots of χ^2/N_{pix} (see Eq. 5) as a function of T for the MAP estimate (C) and the WD+MAP estimate (D).

TABLE 1 Best root-mean-square quality ratios achieved at different values of S/N , using respectively, the RL iterations, the MAP estimate, the WD+MAP estimate, and the WD+RL estimate

Restorations	S/N	Root-mean-square quality ratios (%) at various S/N values			
		4.3	6.2	10.7	13.8
First set, performed with the exact PSF	RL	14.5	23.4	36.9	42.6
	MAP	29.7	31.1	38.4	42.9
	WD+MAP	40.3	41.9	48.1	51.2
	WD+RL	41.3	43.4	49.1	52.2
Second set, performed with the approximate PSF	RL	8.5	14.0	21.2	23.3
	MAP	19.9	19.1	21.1	22.2
	WD+MAP	28.4	26.2	25.5	25.4
	WD+RL	28.5	27.3	26.7	26.7

concern. Their meaning is explained at the end of this section.) We see that Q_2 attains a maximum for a certain temperature T_{opt} (depending on S/N and on the type of estimate considered) that defines the optimal temperature of the restoration in the mean-square sense. In the present example we may say that on the $S/N = 6.2$ image, the MAP and WD+MAP restorations achieve a root-mean-square recovery of 31% and 42%, respectively, when using the correct PSF, whereas they achieve a recovery of 19% and 26%, respectively, when using the approximate PSF. The maximum quality ratios achieved in the other cases are summarized in Table 1. Clearly the use of WD leads to a significant improvement in the quantitative performance of the restoration. In effect, the image is cooled down after WD has been applied, in the sense that the temperature of the optimal MAP restoration is much reduced. This means

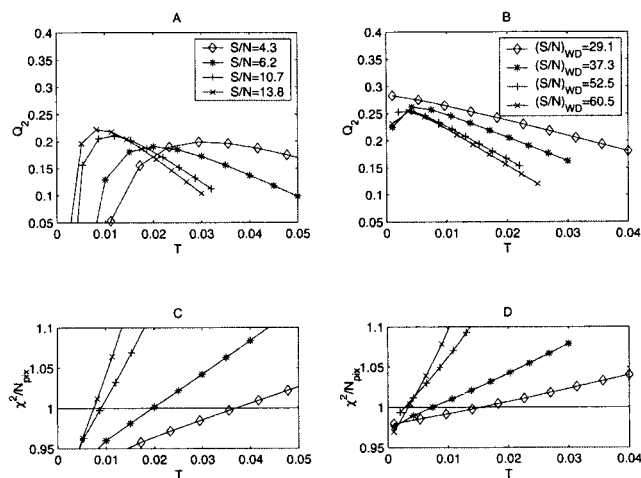


FIGURE 13 Results of MAP and WD+MAP restorations performed on the test image of Fig. 9 (approximate PSF used at $S/N = 4.3, 6.2, 10.7$, and 13.8). (Upper part) Plots of Q_2 as a function of T for the MAP estimate (A) and the WD+MAP estimate (B). (Lower part) Plots of χ^2/N_{pix} (see Eq. 5) as a function of T for the MAP estimate (C) and the WD+MAP estimate (D).

that less regularization is needed in the subsequent deconvolution, which justifies our claim that WD acts as an effective alternative regularization scheme, leading to a less biased restoration. Visually, the WD+MAP estimate also appears more stable (that is, more free of artifacts) than is the MAP estimate, as illustrated in Fig. 14.

Wavelet denoising versus maximum-entropy

The nonregularized Richardson-Lucy algorithm (obtained by setting $T = 0$ in Eq. 3) does not converge to a stable result, but it does lead to a stable improvement during the first few iterations. In particular, after a certain number of iterations, the quality ratio Q_2 defined by Eq. 4 (replacing $f_1(x)$ by the RL estimate) will achieve a maximum value, so that it makes sense to speak of the best RL iteration in the mean-square sense. This is illustrated in Fig. 15, which shows the evolution of Q_2 as a function of the number of iterations for RL iterations (A) and the estimate obtained by applying WD before RL iterations (B; which we refer to as the WD+RL estimate). We see that the WD+RL estimate allows one to achieve a significantly higher Q_2 than RL iterations alone (the difference being quite dramatic at low SNR), and it leads also to a more stable result, in the sense that Q_2 has a much less pronounced maximum after WD.

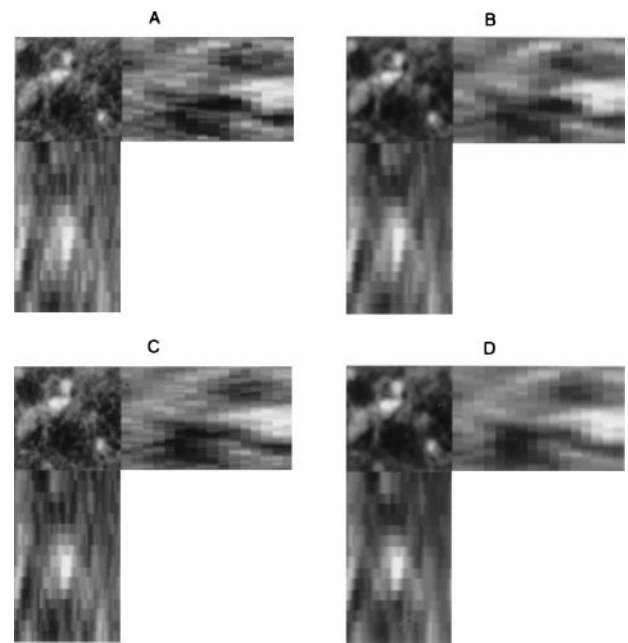


FIGURE 14 Deconvolutions of the test image of Fig. 9 with and without wavelet denoising. (A) Best MAP estimation without applying WD, using the exact PSF ($T_{\text{opt}} \approx 0.015$, $Q_2 \approx 29.7\%$); (B) Best MAP estimate after applying WD, using the exact PSF ($T_{\text{opt}} \approx 0.0042$, $Q_2 \approx 40.3\%$); (C) Best MAP estimation without applying WD, using the approximate PSF ($T_{\text{opt}} \approx 0.02$, $Q_2 \approx 19.9\%$); (D) Best MAP estimate after applying WD, using the approximate PSF ($T_{\text{opt}} \approx 0.0042$, $Q_2 \approx 28.4\%$).

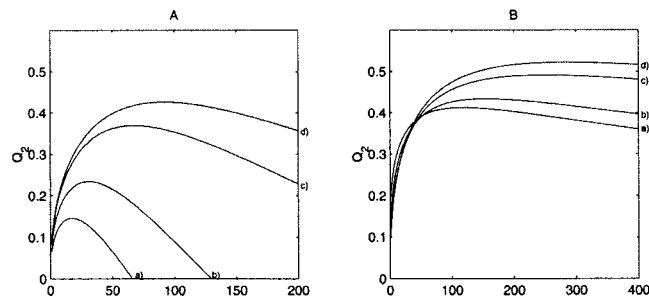


FIGURE 15 Results of RL and WD+RL restorations performed on the test image of Fig. 9 (exact PSF used). (A) Plots of Q_2 as a function of the number of iterations for the RL estimate; (B) Plots of Q_2 as a function of the number of iterations for the WD+RL estimate. The different labels a), b), c), and d) correspond, respectively, to the cases $S/N = 4.3, 6.2, 10.7$, and 13.8 .

It is of interest to compare the mean-square performance of the four estimates defined above (namely, the MAP, WD+MAP, RL, and WD+RL estimates). This comparison is made in Table 1, where the quality ratios achieved by the different estimates in the various cases are regrouped. We see that when applied without denoising, the maximum-entropy constraint enhances the quality of the restoration. More precisely, the quality ratio of the best RL iteration is significantly outperformed by that of the optimal MAP estimate (at least when one is not using a too wrong PSF at a too high SNR). In contrast, the WD+RL estimate appears to achieve a slightly higher quality ratio than does the WD+MAP estimate. The above comparison indicates that what limits the quality of the restoration after WD is not the residual level of noise remaining in the image (note that this noise is no longer Poissonian) but rather other factors to which the maximum-entropy constraint is much less relevant. The level of accuracy of the PSF measurement is probably the most limiting factor, but other factors such as the fineness of pixel sampling and boundary problems may also significantly affect the optimal quality of the restoration. Note that in most of the cases considered, the gain achieved by the WD+RL estimate over the MAP estimate is significant. This clearly supports the claim that WD provides a more effective (and more adaptive (Donoho et al., 1995)) regularization than does the maximum-entropy constraint. Note finally that, even though the WD+RL estimate seems to do a better job than the WD+MAP estimate, the differences are not dramatic, and in fact the two methods lead to nearly equivalent results visually. We find the WD+MAP method more convenient, because it appears to us to be easier in practice to optimize a Bayesian parameter, such as the temperature, than to determine an optimal iteration number.

Choice of the restoration temperature

Some statistical criterion should ideally be used to check the reliability of the restored estimate and to optimize the dif-

ferent restoration parameters in a systematic way. There is a natural χ^2 function for our problem, defined to be the sum of squares of the residuals (Hunt, 1973):

$$\chi^2 = \sum_x (I_n(x) - f * p(x))^2 / \sigma^2, \quad (5)$$

where σ^2 stands for the overall noise variance, defined by $\sigma^2 = 1/N_{\text{pix}} \sum_x \text{Var}(I_n(x))$, which can be accurately estimated from the data (as described in the Appendix). An ideal, error-free estimate for f would correspond to a value of χ^2 equal to the total number N_{pix} of pixels in the image, up to a fluctuation of order $O(\sqrt{N_{\text{pix}}})$. It is thus tempting to use the constraint $\chi^2 \approx N_{\text{pix}}$ as a criterion to choose the value of T . This choice is somewhat naive, as it does not take into account the fact that $f_T(x)$ depends upon $I_n(x)$ in a deterministic way. Although this criterion leads to a significant oversmoothing in the context of least-squares regularization (Galatsanos and Katsaggelos, 1992), in the present case it happens to select a temperature that is not far from optimal. This is illustrated in Figs. 11, C and D, and 12, C and D, showing the behavior of χ^2 as a function of T for our different sets of restorations. Some remarks are in order. First, the maximum-entropy constraint introduces a slight bias on the MAP estimate $f_T(x)$, which must be compensated for when measuring χ^2 : with our normalization, the average intensity $\langle f_T(x) \rangle = N_{\text{pix}}^{-1} \sum_x f_T(x)$ is slightly compressed toward the value 1. Ideally, the deconvolved estimate should have the same average intensity as the original image. It is not difficult to check that this is indeed the case at $T = 0$. In Figs. 11 C and D, and 12, C and D, we therefore computed χ^2 not for $f_T(x)$ but for the adjusted estimate $f_T(x) - \langle f_T(x) \rangle + \langle I_n(x) \rangle$. Second, although the determination of T_{opt} appears to be pretty accurate for the bare MAP estimate, it is less accurate after WD has been applied. The temperature that is selected for the WD+MAP estimate leads to an oversmoothed estimate. However, in practice the above criterion has the advantage of being simple to use, and it provides a useful check of the solution.

CONCLUDING REMARKS

It is clear from the previous sections that the two problems discussed in the Introduction concerning the application of deconvolution to confocal microscopy (inevitable presence of noise in the image and limited knowledge of the PSF) do not prevent one from achieving significant restorations. These two problems can be reduced by combining appropriate processing methods. Bayesian inference offers a powerful approach to deconvolution based on a physically sound model of the microscope's imaging process. In this context, the use of a maximum-entropy constraint provides an effective regularization. Maximum-entropy has its limits, however, both theoretical and practical. Its main justification here is that it leads to simple algorithms and allows one

to actually enhance the quality of the restoration when applied to noisy samples. In principle, such a constraint could be introduced in a less ad hoc way, at the price of using a more complicated Bayesian model of the imaging process (e.g., in the form of Markov random fields (Geman and Geman, 1984; Pryce and Bruce, 1995)). Another (related) possibility of progress, in line with recent work on the statistical physics of image restoration (Nishimori and Wong, 1999; Carlucci and Inoue, 1999), would be to base the restoration not only on the measured intensity $I(x)$, but also on measured correlations between the fluorophore levels of nearby positions inside the specimen. Developing such an image restoration method could be an interesting challenge for image correlation spectroscopy (Petersen et al., 1993).

Wavelet denoising offers an alternative method of regularization for deconvolution. In all the cases that we have considered, this method leads to the restorations of highest quality, in most cases with an important gain in quality ratios over the maximum-entropy MAP restoration. This clearly demonstrates the success of wavelet processing methods in reducing noise problems in 3D microscopy. Of course, wavelets do not provide a cure to an imprecise knowledge of the PSF, but they allow one to impose an adaptive regularization constraint at different scales and different positions in the image, yielding results that would be difficult to achieve with more classic methods at the same computational price. Wavelet analysis is in fact of more general interest to biological microscopy, and it is important to develop multiscale processing methods in a form practical for analyzing complex 3D structures or even 4D processes. In this respect, it would be of interest to design a deconvolution algorithm based entirely on wavelets and practical for the needs of 3D microscopy. Finally, our numerical experiments show clearly the importance of using a PSF that is as accurate as possible when applying deconvolution. One of the main challenges in the conception of deep imaging systems such as confocal and two-photon microscopy is to develop tools allowing a better characterization of the imaging process in situ. This will require not only the conception of new experimental devices, but also the development of efficient adaptive processing methods allowing one to deduce the unknown characteristics of the images from the data. Tools derived from wavelets and the like may play a significant part in the solution of this problem.

APPENDIX: THE DWT AND ITS APPLICATION TO SNR ESTIMATION AND DENOISING

Decimated and translation-invariant DWTs

In the multiresolution setting, the basic ingredients of a DWT are a low-pass filter α and a set of high-pass wavelet filters $\beta^1, \dots, \beta^{n_d}$ satisfying certain orthogonality conditions (there are $n_d = 3$ wavelet filters in two dimensions, 7 in three dimensions, $2^d - 1$ in d dimensions). These

filters are used in a recursive process of convolutions and decimations to produce a series I_1, I_2, \dots of coarse-grained versions of the same digital image I_0 , together with a set of arrays w_1^a, w_2^a, \dots of wavelet coefficients, as follows:

$$I_{j+1,k} = \sum_l \alpha_l I_{j,2k-l}, \quad w_{j+1,k}^a = \sum_l \beta_l^a I_{j,2k-l}, \quad (6)$$

where j labels different scales of resolution in the image, and k and l label the different pixels at given scales. In passing from I_j to I_{j+1} , the resolution is reduced by a factor of 2 and the number of pixels by a factor of 2^d . The finest scale $j = 0$ corresponds to the original image I_0 . In words, we are looking at the image from increasing distances in a geometric (dyadic) progression. The set of arrays $W = \{w_1^a, w_2^a, \dots, w_j^a, I_j\}$ defines the wavelet transform of I_0 at scale level j . For suitably chosen orthogonal filters this transform can be inverted by application of the formula:

$$I_{j,k} = \sum_l I_{j+1,l} \alpha_{2l-k} + \sum_{l,a} w_{j+1,l}^a \beta_{2l-k}^a. \quad (7)$$

(The condition of orthogonality reads $\sum_k \alpha_{2k-l} \alpha_{2k-l'} + \sum_{k,a} \beta_{2k-l}^a \beta_{2k-l'}^a = \delta_{l,l'}$. More generally, one can use for the reconstruction a set of filters $\tilde{\alpha}$ and $\tilde{\beta}^a$ in duality with α and β^a in the sense that $\sum_k \alpha_{2k-l} \tilde{\alpha}_{2k-l'} + \sum_{k,a} \beta_{2k-l}^a \tilde{\beta}_{2k-l'}^a = \delta_{l,l'}$.) This decomposition provides the efficient multiscale analysis of an image evoked in the beginning of the third section: The operators $I_j \rightarrow I_{j+1}$ and $I_j \rightarrow \{w_j^a\}$ play, respectively, the role of the low-pass and high-pass transforms L and H , whereas nonredundancy is ensured by the dyadic decimation process. Due to these decimations, however, the above DWT is not translation invariant, which can be a serious drawback in applications to image restoration. (For instance, the energy sum $\sum_k (w_{j,k}^a)^2$ for a given j changes under a shift of the image.) Translation invariance can be recovered at the price of accepting a certain amount of redundancy in the wavelet coefficients. A convenient procedure is to avoid decimating the data while keeping a dyadic progression of resolution between the successive scales (Mallat, 1999; Coifman and Donoho, 1995).

We used here a discrete analogous to the continuous wavelet transform, based on a discrete scaling function ϕ_{j_0} and discrete wavelet functions ψ_j^a , $j = 0, \dots, j_0$, defined as follows: ϕ_{j_0} is obtained by applying Eq. 7 j_0 times, starting from the array $W_{j_0} = \{w_1^a, w_2^a, \dots, w_{j_0}^a, I_{j_0}\}$ where $w_1^a = 0$, $1 \leq j \leq j_0$, and $I_{j_0,k} = \delta_{k,0}$ (a point-mass placed at the origin of image space at scale level j_0). Similarly, for $j = 1, \dots, j_0$ and $1 \leq a \leq 2^d - 1$, ψ_j^a is the inverse DWT of $W_j^a = \{w_1^a, w_2^a, \dots, w_{j_0}^a, I_{j_0}\}$, where now $I_{j_0} = 0$, $w_{j'}^a = 0$ for $j' \neq j$ or $a' \neq a$, and $w_{j,k}^a = \delta_{k,0}$ (a point-mass placed at the origin of wavelet space at scale level j and polarization a). A translation-invariant DWT (TIDWT) is then obtained by using discrete convolutions of the original image with ϕ_{j_0} and ψ_j^a . Namely, the TIDWT of $I(x)$ is defined by

$$I_{j_0}(x) = \phi_{j_0} * I(x), \quad w_j^a(x) = \psi_j^a * I(x), \quad 1 \leq j \leq j_0, 1 \leq a \leq 2^d - 1, \quad (8)$$

where $f * g(x)$ denotes the cyclic convolution of the two arrays $f(x)$ and $g(x)$. The transformation of Eq. 8 is equivalent (but stated with somewhat different notations) to the TIDWT of Coifman and Donoho (1995). It can be inverted with a reconstruction formula obtained by averaging over the shifts:

$$I(x) = 2^{-j_0 d} \phi_{j_0} * I_{j_0}(x) + \sum_{j=1}^{j_0} \sum_a 2^{-j d} \psi_j^a * w_j^a(x), \quad (9)$$

where we use the notation $\tilde{f}(x) = f(-x)$ for an array $f(x)$. The scaling and wavelet functions we used in our experiments were constructed from 3D-separable filters α and β^a associated with a 1D Daubechies' wavelet for each axis X , Y , and Z .

Estimation of S/N

The SNR of a noisy image $I_n(x) = I_0(x) + n(x)$ (of mean value $I_0(x)$) is defined to be the ratio:

$$\frac{S}{N} = \sqrt{\frac{\langle I_0(x)^2 \rangle}{\langle n(x)^2 \rangle}}, \quad (10)$$

where $\langle \cdot \rangle$ represents spatial average (i.e., average over the pixels of the image). (The SNR is often measured in decibels according to the definition $SNR = 10 \log_{10}(S^2/N^2)$. We stick here to the linear scale defined by Eq. 10, considering the SNR as a true ratio.) Under Poisson statistics, $N = \langle n(x)^2 \rangle$ is related to $S = \langle I_0(x)^2 \rangle$, as the standard deviation of $n(x)$ for a given x is equal to $\sqrt{I_0(x)}$. Perhaps more relevant for imaging purposes is the contrast C of the image, which may be defined as

$$C = \frac{\Delta S}{N} = \sqrt{\frac{\langle (I_0(x) - \langle I_0(x) \rangle)^2 \rangle}{\langle n(x)^2 \rangle}}. \quad (11)$$

In artificial cases, S , N , and C can be computed exactly, but this is of course not the case in practice. When we are given only the image $I_n(x)$, reliable estimates may be obtained with the help of the DWT. To estimate N we assume that the variations of $I_n(x)$ at the finest scales of the image (that is, between two adjacent pixels) are almost entirely caused by noise. Applying an orthogonal DWT to the image and noise, our assumption is equivalent to $w_{1,k}(I_n) \approx w_{1,k}(n)$, where $w_{1,k}(I_n)$ and $w_{1,k}(n)$ denote, respectively, the wavelet coefficients of $I_n(x)$ and $n(x)$ at the finest scale of the transform. This approximation may not be accurate locally throughout the image (in particular it is not adequate in regions where $I_0(x)$ has sharp variations), but it leads to accurate results when taking averages over the pixels k . Because our DWT is an orthogonal transformation, an estimate of N^2 is obtained simply by taking the mean square of the coefficients $w_{1,k}(I_n)$. This estimate is accurate as long as the standard deviation of $n(x)$ does not vary too rapidly with x . A similar estimation could be done with a Fourier transform by taking the mean square of the highest-frequency components, but this would lead to a less accurate estimate, because the number of relevant spatial frequencies is of order $O(N_{\text{pix}}^{1-1/d})$, whereas the number of coefficients $w_{1,k}$ above is $O(N_{\text{pix}})$. To estimate S , we remark that $S^2 = \langle I_0(x)^2 \rangle = \langle I_n(x)^2 \rangle - 2\langle I_0(x), n \rangle - N^2 \approx \langle I_n(x)^2 \rangle - N^2$, up to a relative error (given by the scalar product $\langle I_0(x), n \rangle = \sum_x I_0(x) n(x)$ of order $O(1/\sqrt{N_{\text{pix}}})$). Similarly, we may write $(\Delta S)^2 = \langle (I_0(x) - \langle I_0(x) \rangle)^2 \rangle \approx \langle I_n(x)^2 \rangle - N^2 - \langle I_n(x) \rangle^2$ up to a relative error of order $O(1/\sqrt{N_{\text{pix}}})$. To convince the reader that the above method does not merely give a good idea of what S/N is, but provides a really accurate estimate, we present in Fig. 16 the results that it produces for the test image of Fig. 9. The solid and dashed lines represent the exact values of S/N and N , respectively, whereas the data points are the corresponding estimates computed for a representative set of S/N values and plotted as a function of S/N . The fact that N is nearly equal to S/N reflects the Poissonian nature of the noise statistics. The two lines are nevertheless neatly distinguished by the estimates. The relative error committed on S/N was $\sim 0.5\%$ over the range considered, in accord with the naive expectation, $1/\sqrt{N_{\text{pix}}} \approx 0.003$.

Wavelet denoising

Applying wavelet denoising involves choosing a level of significance in the wavelet domain and thresholding the wavelet coefficients whose magnitudes are below that level. In our experiments we have applied the wavelet shrinkage method (Donoho and Johnstone, 1994). Specifically, this method consists of choosing a threshold parameter $\theta_j^a > 0$ for each wavelet scale j and polarization a and replacing the wavelet coefficients $w_j^a(x)$ of an image by the soft-thresholded estimate

$$\hat{w}_j^a(x|\theta_j^a) = 0 \quad \text{if} \quad |w_j^a(x)| \leq \theta_j^a,$$

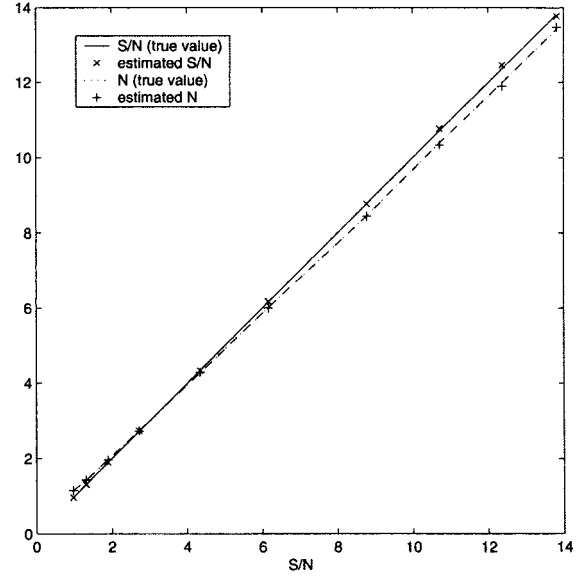


FIGURE 16 Estimation of S/N for the test image of Fig. 9. The solid and dashed lines represent the exact values of S/N and N , respectively, computed for a set of values of S/N in the range $0.97 \leq S/N \leq 13.8$ (corresponding to the decibel interval $(-0.25 \text{ dB}, 23 \text{ dB})$), and line-plotted as a function of S/N . The data points are the estimated values.

$$\hat{w}_j^a(x|\theta_j^a) = w_j^a(x) - \text{sgn}(w_j^a(x))\theta_j^a \quad \text{if} \quad |w_j^a(x)| > \theta_j^a.$$

Soft thresholding can be motivated by adopting a model of the form $w_j^a(x) = s + n$, where the noisy component n follows a Gaussian distribution with mean 0 and standard deviation σ , and the signal s is distributed according to an exponential prior $P(s) = \alpha \exp(-|s|/\alpha)$. The MAP estimate of s under this model is equal to $\hat{w}_j^a(x|\theta_j^a)$ for $\theta_j^a = \sigma^2/\alpha$. Although the motivations of Donoho and Johnstone (1994) for this procedure are different, assuming an exponential prior distribution for the amplitude of wavelet coefficients is not unreasonable empirically (Simoncelli, 1999). When applied to a 3D image, the computation of $\hat{w}_j^a(x|\theta_j^a)$ for, say, scale levels $j = 1, 2, 3$, requires a choice of 21 independent thresholds. Clearly, a systematic and reasonably optimal method for performing this choice is needed. Fortunately, such a method exists in the case where the noise is Gaussian and of uniform intensity throughout the image (Donoho et al., 1995). In brief, this method consists of estimating the mean-square error $\epsilon(\theta)$, or risk, associated with a given threshold θ , which is defined to be the sum $\sum_x (\hat{w}_j^a(I_n)(x|\theta_j^a) - w_j^a(I_0)(x))^2$ averaged over the noise. Knowing the noise distribution allows one to write a general nonbiased estimate of $\epsilon(\theta)$, called the Stein's unbiased risk estimate (SURE) of the problem. Thanks to special properties of the Gaussian distribution, this estimate $\hat{\epsilon}_{\text{SURE}}(\theta)$ takes a very simple form in the present case:

$$\hat{\epsilon}_{\text{SURE}}(\theta) = \sigma^2 N_{\text{pix}} - 2\sigma^2 N_{\leq \theta} + \sum_x \min(w_j^a(x), \theta)^2, \quad (12)$$

where $N_{\leq \theta}$ is the number of pixels x such that $|w_j^a(x)| \leq \theta$, and σ^2 is the noise variance. Minimization of $\hat{\epsilon}_{\text{SURE}}(\theta)$ over θ can be done very efficiently and leads to a choice of θ_j^a , which becomes optimal in the limit of large images (i.e., as $N_{\text{pix}} \rightarrow \infty$) (Donoho et al., 1995).

Under Poisson statistics the noise variance is directly proportional to the image's mean $I_0(x)$ and is therefore not uniform. We expect, however, the SURE method to perform reasonably well when applied on a relatively uniform region of the image. This is pretty well confirmed from experi-

ments performed on small test images, as shown in Fig. 10. In these experiments we applied the wavelet shrinkage method to the scale levels $j = 1, 2, 3$ of the TIDWT of $I_n(x)$. In another set of experiments (not shown), we also found that even for the relatively small image size considered here, the SURE thresholds are pretty close to optimal. We point out that, due to the finite sampling resolution of the image, there is a transition value $(S/N)_c$ below which the denoised estimate becomes depinned from the signal, in the sense that its correlation with the non-noisy image that we try to recover falls rapidly from nearly 1 to nearly 0 across $(S/N)_c$. The gain $(S/N)_{WD}/(S/N)$ corresponds to a genuine improvement of SNR only in the region where S/N is reasonably larger than $(S/N)_c$. In the present example $(S/N)_c \approx 0.5$ (~ -6 dB). For S/N just below $(S/N)_c$ the gain dropped, and it stabilized around 6.7 for extremely small S/N . This limit gain reflects, however, a trivial gain of information: the denoised estimate is essentially uncorrelated with the true signal, but it is closer to the true signal because it is closer to 0 (actually a higher gain is achieved by the uniform estimate 0).

When considering a large image with possibly significant variations of brightness, we may still apply the wavelet shrinkage procedure, using the time space-varying thresholds θ_j^S . In our experiments we applied a set of thresholds obtained by standard interpolation from local SURE thresholds calculated for an ensemble of small patches covering the image. (Typically, we used patches of size $64 \times 64 \times 16$ or $64 \times 64 \times 32$.)

We thank the referees for many comments that helped to improve the paper. We thank also Prof. Joseph Bruton for reading the manuscript with great care and pointing out to us many unclear sentences and misspellings. This work was supported by grants from the Swedish Medical Research Council (09888), the Swedish Council for Work Life Research (1999–0512), the Foundation Tysta Skolan, the Swedish Institute, the Knut and Alice Wallenberg Foundation, and Institut National de la Santé et de la Recherche Médicale (INSERM). Parts of this project were carried out within the framework of the European Union research program (BIO4-CT98–0408).

REFERENCES

- Bertero, M., P. Boccacci, G. Brakenhoff, F. Malfanti, and H. van der Voort. 1990. Three-dimensional image restoration and super-resolution in fluorescence confocal microscopy. *J. Microsc.* 157:3–20.
- Candès, E., and D. Donoho. 2000. Recovering edges in ill-posed inverse problems: optimality of curvelet frames. Technical report. Stanford University, Stanford, CA.
- Carlucci, D., and J.-I. Inoue. 1999. Image restoration using the chiral potts spin-glass. *Phys. Rev. E*. 60:2547–2553.
- Coifman, R., and D. Donoho. 1995. Translation-invariant de-noising. In *Wavelets in Statistics: Lecture Notes in Statistics*. A. Antoniadis and G. Oppenheim, editors. Springer-Verlag, Berlin. 125–150.
- Csiszár, I. 1991. Why least-squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Ann. Stat.* 19: 2033–2066.
- Dempster, A., N. Laird, and D. Rubin. 1977. Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. B*. 39:1–37.
- Donoho, D. L. 1995. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmonic Anal.* 2:101–126.
- Donoho, D. L., and I. M. Johnstone. 1994. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*. 81:425–455.
- Donoho, D. L., I. M. Johnstone, and M. Iain. 1995. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Stat. Assoc.* 90:1200–1224.
- Fligge, M., and S. Solanski. 1997. Noise reduction in astronomical spectra using wavelet packets. *Astronomy Astrophys. Suppl. Ser.* 124:579–587.
- Flock, Å., B. Flock, A. Fridberger, E. Scarfone, and M. Ulfendahl. 1999. Supporting cells contribute to control of hearing sensitivity. *J. Neurosci.* 19(11):4498–4507.
- Flock, Å., E. Scarfone, and M. Ulfendahl. 1998. Vital staining of the intact hearing organ: visualization of cellular structure with confocal microscopy. *Neuroscience*. 83:215–228.
- Galatsanos, N., and A. Katsaggelos. 1992. Methods for choosing the regularization parameter and estimating noise variance in image restoration and their relation. *IEEE Trans. Im. Proc.* 1:322–336.
- Geman, S., and D. Geman. 1984. Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. *IEEE Trans. on Patt. Anal. and Mach. Intell.* 6:721–741.
- Gibson, S., and F. Lanni. 1991. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *J. Opt. Soc. Am. A*. 8(10):1601–1613.
- Gull, S., and G. Daniell. 1978. Image reconstruction from incomplete and noisy data. *Nature*. 272:686–690.
- Hell, S., G. Reiner, C. Cremer, and E. Stelzer. 1993. Aberrations in confocal fluorescence microscopy induced by mismatches in refractive index. *J. Microsc.* 169:391–405.
- Holmes, T. J. 1988. Maximum-likelihood image restoration adapted for noncoherent optical imaging. *J. Opt. Soc. Am.* A5:666–673.
- Hunt, B. R. 1973. The application of constrained least squares estimation to image restoration by digital computer. *IEEE Trans. Comput.* 22: 805–812.
- Jansson, P., R. Hunt, and E. Plyler. 1970. Resolution enhancement of spectra. *J. Opt. Soc. Am.* 60:596–599.
- Le Calvez, S., and M. Ulfendahl. 2000. An in vitro preparation to access cellular and neuronal components in the mouse inner ear. *J. Neurocytol.* 29:645–653.
- Lucy, L. B. 1974. An iterative technique for the rectification of observed distributions. *Astronomical J.* 79:745–754.
- Mallat, S. 1999. *A Wavelet Tour of Image Processing*. Academic Press, New York.
- Marroquin, J., S. Mitter, and T. Poggio. 1987. Probabilistic solution of ill-posed problems in computational vision. *J. Am. Stat. Assoc.* 82: 76–89.
- Miller, K. 1970. Least squares methods for ill-posed problems with a prescribed bound. *SIAM J. Math. Anal.* 1:52–74.
- Nishimori, H., and K. Wong. 1999. Statistical mechanics of image restoration and error correcting codes. *Phys. Rev. E*. 60:132.
- Nowak, R., and E. Kolaczyk. 2000. A statistical multiscale framework for Poisson inverse problems. *IEEE Trans. Info. Theory*. 46:1811–1825.
- Pawley, J. 1995. *Handbook of Biological Confocal Microscopy*. Plenum Press, New York.
- Petersen, N., P. Höddelius, P. Wiseman, O. Seger, and K. Magnusson. 1993. Quantitation of membrane receptors distribution by image correlation spectroscopy: concept and application. *Biophys. J.* 49:809–815.
- Pryce, J., and A. Bruce. 1995. Statistical mechanics of image restoration. *J. Phys. A Math. Gen.* 28:511–532.
- Richardson, W. H. 1972. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* 62:55–59.
- Shepp, L. A., and Y. Vardi. 1982. Maximum-likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imag.* 1:113–122.
- Simoncelli, E. 1999. Modeling the joint statistics of images in the wavelet domain. In *Proc. SPIE Annu. Meet., 44th, Denver, CO*. 3813:188–195.
- Starck, J. L., and A. Bijaoui. 1994. Filtering and deconvolution by the wavelet transform. *Signal Processing*. 35:195–211.
- Strömberg, J.-O. 1997. Wavelets and application to medical images. *Proc. SSAB' 97: Swedish Symp. Image Anal., Stockholm*. 110–113.
- Tikhonov, A., and V. Arsenin. 1977. *Solutions of Ill-Posed Problems*. Winston/Wiley, Washington.
- Ulfendahl, M., Å. Flock, and S. Khanna. 1989. A temporal bone preparation for the study of cochlear micromechanics at the cellular level. *Hearing Res.* 40:55–64.
- Ulfendahl, M., E. Scarfone, Å. Flock, S. Le Calvez, and P. Conradi. 2000. Perilymphatic fluid compartments and intercellular spaces of the inner ear and organ of corti. *NeuroImage*. 12:307–313.