

# Improved Bayesian image denoising based on wavelets with applications to electron microscopy

C.O.S. Sorzano<sup>a,b,\*</sup>, E. Ortiz<sup>a</sup>, M. López<sup>c</sup>, J. Rodrigo<sup>d</sup>

<sup>a</sup>*Dept. Sistemas Electrónicos y de Telecomunicación, Escuela Politécnica Superior, Univ. San Pablo-CEU, Urb. Montepríncipe s/n, Boadilla del Monte, 28668 Madrid, Spain*

<sup>b</sup>*Biocomputing Unit, National Center of Biotechnology (CSIC), Campus Univ. Autónoma s/n, 28049 Madrid, Spain*

<sup>c</sup>*Dept. Matemática e Informática Aplicadas a la Ingeniería Civil, E.T.S. Ingenieros de Caminos, Univ. Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain*

<sup>d</sup>*Dept. Matemática Aplicada, E.T.S. Ingeniería, Univ. Pontificia Comillas, c/Alberto Aguilera, 23, 28015 Madrid, Spain*

Received 14 June 2004; received in revised form 1 December 2005; accepted 1 December 2005

---

## Abstract

In this work we discuss an improvement of the image-denoising wavelet-based method presented by Bijaoui [Wavelets, Gaussian mixtures and Wiener filtering, *Signal Process.* 82 (2002) 709–712]. We show that the parameter estimation step can be replaced by a constrained nonlinear optimization. We propose three different methods to estimate the parameters. As in Bijaoui's original article, two of them deal with white noise. We show that the resulting algorithms improve the one originally proposed. Our third method extends the applicability of the denoising algorithm to colored noise. We test our algorithms with images simulating electron microscopy (EM) conditions as well as experimental EM images.

© 2006 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

*Keywords:* Image denoising; Bayesian filtering; Wavelets

---

## 1. Introduction

Many are the papers addressing the problem of image denoising using wavelets [1]. In brief, all these algorithms first perform the wavelet transform of the image to denoise, then apply some filter to the wavelet coefficients, and finally take the inverse wavelet transform to restore the denoised image. Most popular wavelet-filtering algorithms are based on thresholding [2], Wiener filtering [3] and Bayesian filtering [4–7]. In fact, Bayesian filtering in the wavelet space has become the standard in the field and the work of Portilla et al. [7] represents the current state of the art of the algorithms employed.

Bijaoui [1] introduced a Bayesian approach to image denoising in wavelet space shown to be superior to a number of previous thresholding or Wiener-filtering algorithms.

However, as is shown in this paper, the parameter estimation step of this algorithm can be substantially improved. As an alternative, we propose three different methods to estimate the noise and signal parameters, keeping the Bayesian denoising step proposed by Bijaoui [1]. Each method is based on different assumptions and, ultimately, they solve a linear equation system in a least-squares sense subject to a number of reasonable constraints.

We tested our algorithms on a set of simulated as well as experimental EM images as used in single-particle structural studies of macromolecular complexes [8]. The final goal of this kind of studies is to elucidate the three-dimensional structure of the specimen electron density combining different projection images taken with an electron microscope. One of the advantages of this technique is the wide range of specimen sizes that can be studied: from cellular organelles like the ribosome, to viruses, protein complexes, or individual proteins. Another advantage, particularly when dealing with proteins or protein complexes, is that the

---

\* Corresponding author. Tel.: +34 91 372 4033; fax: +34 91 372 4049.

E-mail address: [coss.eps@ceu.es](mailto:coss.eps@ceu.es) (C.O.S. Sorzano).

technique does not change the structural conformation, because macromolecules are not crystallized. This makes the technique suitable for studies where the quaternary structure of protein complexes has to be elucidated. The resolution that can be obtained depends on the quality, quantity and angular coverage of the projection images.

The main drawbacks of EM images are their extremely low signal-to-noise ratio (below  $-3$  dB) and the complicated nature of the electron microscope transfer function [9]. This function is referred to as contrast transfer function (CTF) and it must be estimated directly from the micrographs [10]. The shape of the CTF in Fourier space can be described by a damped sinusoid. However, in a rough approximation it can be modelled as a low-pass filter. The filtered nature of these images make Bijaoui's approach unapplicable. However, we show that the method can be generalized to address this kind of images.

Image denoising techniques can be applied in two ways in EM: first, if the image signal components are not modified, then denoising techniques can be used before reconstructing the three-dimensional structure of the macromolecule to improve the quality of the reconstruction; second, if the image signal components are modified but the result enhances most of the main features of the particle, then they can be used to facilitate intermediate processes like bidimensional alignments, automatic particle picking, etc.

This paper is organized as follows. First, the image denoising procedure based on a Bayesian approach in the wavelet space is introduced in Section 2. Then, we present our proposed modifications to the parameter estimation step in Section 3. Results obtained with simulated as well as experimental data are shown in Section 5. Finally, we discuss the results and conclude.

## 2. The filtering procedure

The filtering technique used in this article is based on the Bayesian approach proposed by Bijaoui [1]. Let us assume that the image formation model is of the form

$$y = x + n, \quad (1)$$

where  $x$  is the ideal image,  $n$  is random independent noise, and  $y$  is the measured image. We assume that the noise is white and normally distributed. Thus, its probability density function (PDF) is given by

$$p(n) = G(n, N) = \frac{1}{\sqrt{2\pi N}} e^{-n^2/2N}, \quad (2)$$

where  $N$  is the noise variance. Solving the image formation model for  $n$  it can be seen that the conditional PDF of  $y$  given  $x$  is

$$p(y|x) = p(n) = G(y - x, N). \quad (3)$$

It is assumed that the PDF of the ideal image can be expressed as a mixture of zero-mean Gaussians

$$p(x) = \sum_i \alpha_i G(x, S_i). \quad (4)$$

Under these assumptions it can be proved that the PDF of the measured image is given by the convolution of both PDFs

$$p(y) = p(x) \star p(n) = \sum_i \alpha_i G(x, S_i + N). \quad (5)$$

Then, the denoising problem is stated as the Bayesian problem of estimating  $x$  from  $y$ . The a posteriori PDF is given by

$$p(x|y) = \frac{\sum_i \alpha_i G(x, S_i) G(y - x, N)}{\sum_i \alpha_i G(y, S_i + N)}. \quad (6)$$

The filtering is done by taking the a posteriori expectation of  $x$

$$E\{x|y\} = y \frac{\sum_i \alpha_i (S_i/S_i + N) G(y, S_i + N)}{\sum_i \alpha_i G(y, S_i + N)} = y w(y). \quad (7)$$

The relationship of this filter with classical Wiener filtering is given in the article by Bijaoui [1].

The previous filter is applied independently to each of the scales of the wavelet transform of the measured image, i.e., the variable  $x$  is formed by all those wavelet coefficients belonging to the same scale. Therefore, each scale has its own  $\alpha_i$  and  $S_i$  parameters while the noise power,  $N$ , is common to all the scales since the noise is assumed to be white.

The problem now is to estimate the  $\alpha_i$ ,  $S_i$  and  $N$  parameters from the measured images. Bijaoui [1] assumes that there is no signal at the finest scale. Therefore, all the energy at that scale is coming from the noise term. Under this assumption the noise power  $N$  can be estimated from that scale and later used for the rest of the scales. Bijaoui [1] proposes a robust estimate of  $N$  at the finest scale based on a  $k - \sigma$  clipping strategy.

For the rest of the scales, the original paper affirms that, in practical terms, only two Gaussians are necessary to model the PDF of the measured wavelet coefficients at a given scale

$$p(y) = (1 - a)G(y, N) + aG(y, S + N), \quad (8)$$

where  $a$  and  $S$  are estimated at each scale through the 2nd ( $M_2$ ) and fourth order ( $M_4$ ) moments of the measured wavelet coefficients at that scale

$$S = \frac{M_4/3 - N^2}{M_2 - N},$$

$$a = \frac{(M_2 - N)^2}{M_4/3 - N^2}. \quad (9)$$

In order to guarantee the convex nature of the combination given by Eq. (8)  $a$  is set to 0 if  $M_2 - N < 0$  or if

$M_4/3 - N^2 < 0$ . Accordingly,  $a$  is set to 1 if  $a > 1$  and, in this case,  $S$  is estimated only from the variance  $M_2$ .

### 3. Improvement of the parameter estimation

The main contribution of our study refers to the estimation of the parameters  $S$ ,  $a$  and  $N$ . We provide three new estimation procedures, each one with different hypothesis and applicable to different fields. In the first one, we simultaneously solve for all the parameters using a rationale similar to the one used by Bijaoui [1]. The first method assumes that the wavelet coefficients follow a Gaussian distribution and that the noise is white and independent from the signal. In the second one, we solve for the same denoising problem (removal of white noise) following a different approach based on the power decomposition of the measured image. This second method assumes that the noise is white and independent from the signal. Finally, we extend the power decomposition procedure to the case of non-white noise assuming that the noise is independent from the signal.

#### 3.1. Method 1

In this method we solve simultaneously for all the noise and signal parameters in the problem. Our changes also affect the “practical” PDF given by Eq. (8). For the kind of images addressed in this work a single Gaussian suffices to model the ideal wavelet coefficients,  $x$ , at a given scale (this point will be explicitly checked in our experiments). In this way, our “practical” PDF becomes

$$p(y) = G(y, S + N). \tag{10}$$

This model has an advantage over the model in Eq. (8) since the estimates given by Eq. (9) may become highly unstable under certain circumstances. Let us assume that there is no signal energy at the first (finest) and second scales. Under these circumstances, the noise power is correctly estimated from the first scale. However, when estimating the signal power at the second scale we are using the second and fourth order moments of the Gaussian defining the noise, that should be the same as for the first scale. Therefore,  $M_2$  and  $M_4$  for the second scale take the value  $N$  and  $3N^2$ , respectively, (for a Gaussian PDF it can be proved that  $M_4 = 3M_2^2$ ). Plugging these values into Eq. (9) we have 0 over 0, and consequently the  $S$  and  $a$  parameters cannot be determined. In practical terms, 0 over 0 is never obtained due to the random nature of the estimates of the  $M_2$  and  $M_4$  moments. However, the resulting  $S$  and  $a$  parameters may be highly unstable.

Using Eq. (10) an alternative equation system can be posed. As in the original article, we assume that there is no significant signal at the finest scale. However, we do not estimate the noise power only from the first scale ( $s = 1$ ), and subsequently all the rest of parameters, but we estimate all

the parameters simultaneously in a least squares sense constrained by some reasonable hypothesis.

Concisely, given the PDF model in Eq. (10), the following equation must hold for all scales  $s = 1, 2, \dots$

$$M_{2s} = S_s + N, \tag{11}$$

where  $M_{2s}$  is the variance of the measured wavelet coefficients at the scale  $s$  and  $S_s$  is the variance of the ideal wavelet coefficients at that scale. Furthermore, we know that all variances must be positive ( $N \geq 0$  and  $S_s \geq 0$ ), and for the kind of images employed in this article  $S_s \leq S_{s+1}$  (i.e., most of the energy is concentrated at low frequencies). Finally, if reasonable lower and upper limits for the signal-to-noise ratio ( $SNR_l$  and  $SNR_h$ ) are provided, then it must hold that

$$\frac{I}{1 + SNR_h} \leq N \leq \frac{I}{1 + SNR_l}, \tag{12}$$

where we have made use of the fact that the total image power,  $I$ , must be  $I = S_T + N$  ( $S_T$  is an estimate of the total signal power). In case that no clue is available about the SNR ( $SNR_l = 0$  and  $SNR_h = \infty$ ) the bounds of the noise power become  $0 \leq N \leq I$ .

To solve simultaneously for all the parameters we propose to minimize  $\|C\mathbf{x} - \mathbf{M}_2\|^2$  subject to  $A\mathbf{x} \leq \mathbf{b}$ , where

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 1 & 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} N \\ S_2 \\ S_3 \\ S_4 \\ \dots \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} M_{21} \\ M_{22} \\ M_{23} \\ M_{24} \\ \dots \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 1 & -1 & 0 & \dots \\ 0 & 0 & 1 & -1 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & 0 & 0 & \dots \\ 0 & -1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & \dots \\ 0 & 0 & 0 & -1 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \end{pmatrix},$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ \dots \\ I \\ \frac{1 + SNR_h}{I} \\ \frac{I}{1 + SNR_l} \end{pmatrix}.$$

The main advantages of the proposed method with respect to the original paper are its simultaneous nature (allowing for a higher stability) and the possibility of constraining the parameters with a priori information (SNR and the constrain that for a certain kind of images  $S_s \leq S_{s+1}$ ).

### 3.2. Method 2

The second method is based on the power decomposition of the measured image. If the image formation model is  $y = x + n$ , and  $x$  and  $n$  are independent variables, then the measured image power  $I = P_y$  is the sum of the corresponding powers  $P_x$  and  $P_n$ . Since the wavelet decomposition is orthogonal, this fact is also true at any scale of the wavelet decomposition. Furthermore, due to the orthonormality of our wavelet transform  $I = P_{wy}$ , where  $P_{wy}$  is the power of the wavelet transform of the image  $y$ .

In this way, the power of the wavelet transform of  $y$  at each scale  $s$ ,  $P_s$ , can be decomposed as the sum of two components: one coming from noise and another one coming from the signal

$$P_s = (N + S_s)n_s, \tag{13}$$

where  $n_s$  is the number of wavelet coefficients at scale  $s$ . Again, we assume that there is no significant signal at the finest scale ( $S_1 = 0$ ). The power decomposition must be constrained in such a way that the sum of the power at all scales yields the overall image power

$$I = \sum_s P_s. \tag{14}$$

Eq. (12) can still be used to incorporate our a priori information about noise.

This second procedure also solves simultaneously for all the noise and signal parameters involved minimizing  $\|C\mathbf{x} - \mathbf{P}\|^2$  subject to  $\mathbf{Ax} \leq \mathbf{b}$  and  $A_{eq}\mathbf{x} = \mathbf{b}_{eq}$ , where

$$C = \begin{pmatrix} n_1 & 0 & 0 & 0 & \dots \\ n_2 & n_2 & 0 & 0 & \dots \\ n_3 & 0 & n_3 & 0 & \dots \\ n_4 & 0 & 0 & n_4 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} N \\ S_2 \\ S_3 \\ S_4 \\ \dots \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ \dots \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 1 & -1 & 0 & \dots \\ 0 & 0 & 1 & -1 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & 0 & 0 & \dots \\ 0 & -1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & \dots \\ 0 & 0 & 0 & -1 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ \dots \\ \frac{I}{1 + SNR_h} \\ \frac{I}{1 + SNR_l} \end{pmatrix},$$

$$A_{eq} = \left( \sum_s n_s \ n_2 \ n_3 \ n_4 \ \dots \right) \quad \text{and} \quad \mathbf{b}_{eq} = (I).$$

This second method is based on a weaker assumption on the image formation model: noise is independent from signal. A drawback of the two procedures presented above is that they assume that noise is white (i.e., it has the same power at all frequencies) and that there is no significant signal at the finest wavelet scale. This precludes the application of this denoising method to filtered images. Next procedure will partially overcome this problem.

### 3.3. Method 3

If images are low-pass filtered, for instance, then the noise power measured at the finest scale does not correspond to the true noise power present at the rest of the scales. This is a basic assumption in Methods 1 and 2. Furthermore, in those methods it was assumed that there was no significant signal at the finest scale, which may not always be the case. In this last procedure, we provide a framework in which both situations can be handled easily. It is an extension of Method 2 and, therefore, it is based on the decomposition of the image power in different components due to different sources. The main departure from the previous method is that we allow for different noise powers  $N_s$  at each scale and that  $S_1$  is not necessarily null. Therefore, the power decomposition at each scale is now

$$P_s = (N_s + S_s)n_s. \tag{15}$$

Eq. (14) still must hold because signal and noise are again independent. The SNR constraints adopt a different expression

$$SNR_l \leq \frac{\sum_s n_s S_s}{\sum_s n_s N_s} \leq SNR_h. \tag{16}$$

Additionally, we impose the constraint  $N_s \leq N_{s+1}$ . This constraint assumes that the imaging system attenuates high frequencies more than low frequencies.

Summarizing, the cost function to minimize is still  $\|C\mathbf{x} - \mathbf{P}\|^2$  subject to  $A\mathbf{x} \leq \mathbf{b}$  and  $A_{eq}\mathbf{x} = \mathbf{b}_{eq}$ , where now

and the input vector  $\mathbf{x}$  is assumed to follow a multivariate normal distribution with zero mean and covariance matrix  $\Sigma_x$ , i.e.,  $p(\mathbf{x}) = G(\mathbf{0}, \Sigma_x)$ . Under these assumptions it has been proven [11, Chapter 46] that the filter  $W$  to optimally recover  $\tilde{\mathbf{x}} = W\mathbf{y}$  from  $\mathbf{y}$  in a maximum a posteriori

$$C = \begin{pmatrix} n_1 & 0 & 0 & 0 & \dots & n_1 & 0 & 0 & 0 & \dots \\ 0 & n_2 & 0 & 0 & \dots & 0 & n_2 & 0 & 0 & \dots \\ 0 & 0 & n_3 & 0 & \dots & 0 & 0 & n_3 & 0 & \dots \\ 0 & 0 & 0 & n_4 & \dots & 0 & 0 & 0 & n_4 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \mathbf{x} = \begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \\ \dots \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ \dots \end{pmatrix}, \mathbf{P} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ \dots \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & -1 & \dots & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -1 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & -1 & \dots \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & -1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & -1 & \dots & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & -1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & -1 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & -1 & \dots \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots \\ -SNR_h n_1 & -SNR_h n_2 & -SNR_h n_3 & -SNR_h n_4 & \dots & n_1 & n_2 & n_3 & n_4 & \dots \\ SNR_l n_1 & SNR_l n_2 & SNR_l n_3 & SNR_l n_4 & \dots & -n_1 & -n_2 & -n_3 & -n_4 & \dots \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \end{pmatrix},$$

$$A_{eq} = \left( \sum_s n_s \ n_2 \ n_3 \ n_4 \ \dots \right) \text{ and } \mathbf{b}_{eq} = (P_{wy}).$$

#### 4. Relationship to Wiener filtering

In our model, it has been assumed that the PDF of the signal coefficients can be represented with a single Gaussian. In this case, the filter applied falls back to a classical Wiener filter. The hypothesis of the Wiener filter is that the relationship between the original image  $\mathbf{x}$  ( $\mathbf{x}$  is a vector with all pixel values ordered lexicographically) and the acquired image  $\mathbf{y}$  is

$$\mathbf{y} = H\mathbf{x} + \mathbf{n}, \tag{17}$$

where  $H$  is a blurring matrix and  $\mathbf{n}$  is a random noise vector.  $\mathbf{n}$  is assumed to follow a multivariate normal distribution with zero mean and covariance matrix  $\sigma^2 I$ , i.e.,  $p(\mathbf{n}) = G(\mathbf{0}, \sigma^2 I)$

sense is

$$W = \Sigma_x H^t (\sigma^2 I + H \Sigma_x H^t)^{-1}, \tag{18}$$

where  $H^t$  denotes the transpose of  $H$ .

Taking into account that we are considering one-dimensional vectors, for us  $H = I$ , and that in our notation  $\Sigma_x = S$  and  $\sigma^2 = N$  this filter boils down to

$$W = \frac{S}{N + S} = \frac{1}{1 + (1/SNR)}. \tag{19}$$

In other words, our Bayesian filtering is the same as a Wiener filter and our contribution lays solely on how to estimate the filter parameters  $N$  and  $S$  at each wavelet scale.

## 5. Results

### 5.1. Simulated data

In order to test the efficacy of the newly proposed parameter estimation methods we denoised 600 projection images (of size  $64 \times 64$ ) of the Protein Data Bank [12] model of the bacteriorhodopsin [13] (PDB entry: 1BRD, see Fig. 1). These projection images were computed as line integrals of a voxelized volume created from the atom description provided by the PDB model. The Xmipp package [14] was used to create these projections. In our first experiment we tested the performance of Methods 1 and 2 with respect to Bijaoui's. We added white Gaussian noise up to a SNR of  $-10$  dB. These parameters yield similar images to those obtained in electron cryomicroscopy [9]. Bijaoui's method as well as Methods 1 and 2 were applied to an orthonormal wavelet decomposition (with Daubechies 12 as wavelet function [15]) of the input images. We applied the denoising procedures in the first four scales ( $s = 1, 2, 3, 4$ ) since further scales had too few wavelet coefficients to make reliable estimates of the statistical moments. We used  $SNR_l = -\infty$ (dB) and  $SNR_h = 0$ (dB) as minimization constraints.

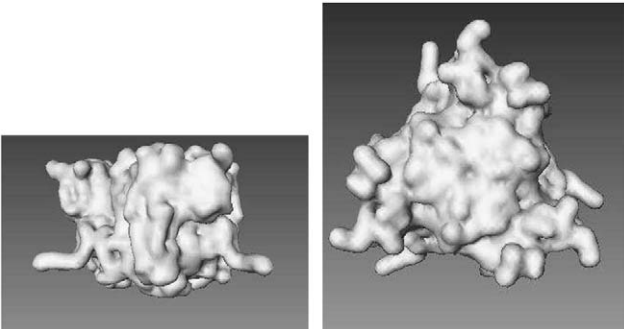


Fig. 1. Side and top view of the isosurface of the bacteriorhodopsin.

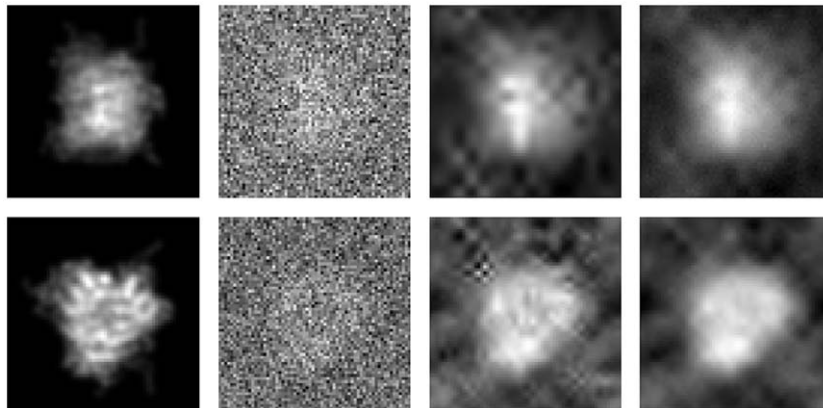


Fig. 2. Each row shows a different example of the results for cryomicroscopy conditions of the two denoising procedures (Bijaoui's and Method 1). From left to right: original image, noisy image, image denoised using original Bijaoui's method, image denoised using Method 1.

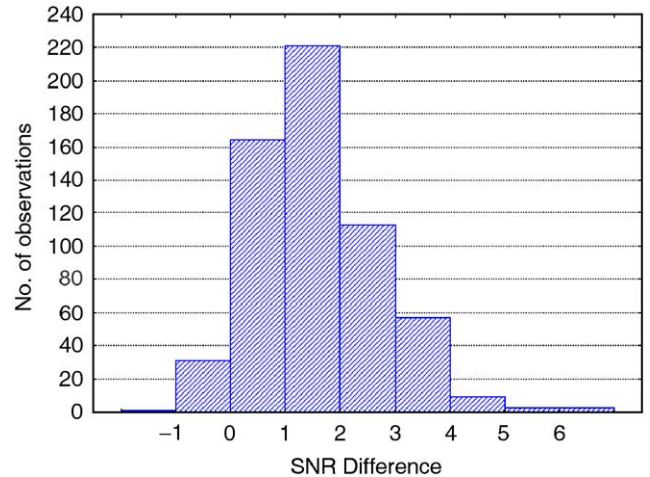


Fig. 3. Histogram of the SNR difference between the proposed Method 1 and the one originally presented by Bijaoui for the images simulating cryomicroscopy. Positive differences favors our modification.

Before applying Method 1 we contrasted the hypothesis that the wavelet coefficients at each scale were coming from a normal distribution by performing the Geary's test [16]. The null hypothesis could not be rejected with a confidence of 99% and, therefore, the wavelet coefficients came from a single Gaussian as assumed by Method 1.

Fig. 2 shows examples of the resulting images after applying Bijaoui's method and Method 1. Fig. 3 shows the histogram of the difference between our output SNR (Method 1) and Bijaoui's output SNR. The average SNR of the denoised image using the original method (Bijaoui) was  $7.06 \pm 1.12$ (dB), while the average SNR of the denoised image using the newly proposed method (Method 1) was  $8.60 \pm 0.76$ (dB). No significant difference (with a 99% confidence) was detected between the performances of Methods 1 and 2 in this experiment.

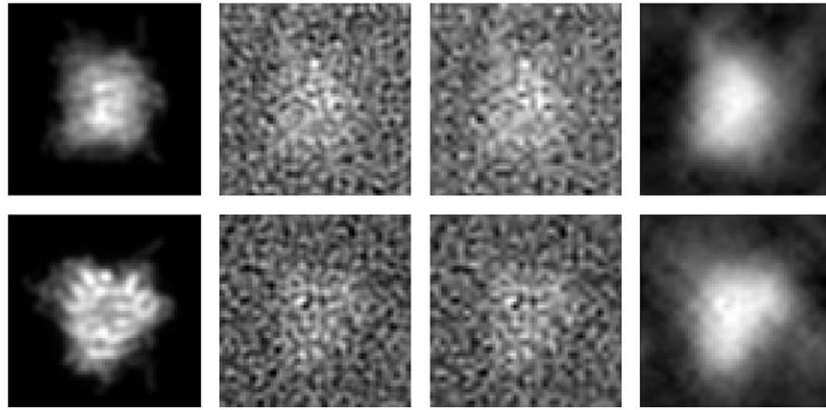


Fig. 4. Each row shows a different example of the results for negative staining conditions of the two denoising procedures (Bijaoui's and Method 3). From left to right: original image, noisy image, image denoised using original Bijaoui method, image denoised using Method 3.

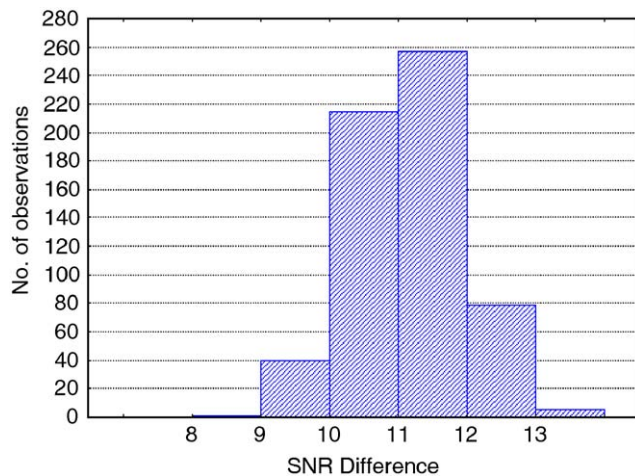


Fig. 5. Histogram of the SNR difference between the proposed method (Method 3) and the one originally presented by Bijaoui for the images simulating negative staining. Positive differences favors our modification.

In our second simulated experiment we tested the efficacy of the third estimation procedure. We simulated the conditions obtained in negative-staining microscopy [9] by adding white Gaussian noise (up to a SNR of  $-10$  dB) and then low-pass filtering the noisy image with a cutoff frequency of  $1/15 \text{ \AA}^{-1}$ . We used  $SNR_l = -30$  (dB) and  $SNR_h = -5$  (dB) as minimization constraints. Notice that in this case, since we stopped denoising at  $s = 4$ , the SNR must represent our knowledge about the SNR in the first four scales.

Fig. 4 shows examples of the resulting images after applying Bijaoui's method and Method 3. Fig. 5 shows the histogram of the difference between our output SNR (Method 3) and Bijaoui's output SNR. The average SNR of the denoised image using the original method (Bijaoui) was  $-1.66 \pm 0.11$  (dB), while the average SNR of the denoised image using the newly proposed method (Method 3) was  $9.49 \pm 0.79$  (dB).

## 5.2. Experimental data

Finally, we applied Method 3 to denoise negative staining micrographs of the DnaB protein [17]. These images were taken with a Philips CM120 electron microscope equipped with a Gatan cryostage. Defocus is between  $1.0 \mu\text{m}$  and  $1.5 \mu\text{m}$ . The pixel size is  $3.6 \text{ \AA}$ . Fig. 6 shows an example of a  $512 \times 512$  piece of such images and its corresponding denoised image. It can be seen that particle projections are much more noticeable in the denoised image.

## 6. Discussion and conclusion

The Bayesian approach proposed by Bijaoui [1] was proven to compare favorably to other wavelet-thresholding denoising techniques. The procedures introduced in this article (Methods 1 and 2) improved the performance of Bijaoui's methodology by solving simultaneously for all the signal and noise parameters involved. In the first simulated experiment carried out, no significant difference was observed between Methods 1 and 2. This is an interesting result since the rationales of both methods are different. Method 1 is based on the analysis of the PDF of the measured image, while Method 2 is based on its power decomposition. Besides, our parameter search can be optionally constrained by a priori knowledge about the SNR or about the nature of the images (in our case, energy tends to concentrate in low-frequency components). The application of the modified methodology to images simulating cryomicroscopy conditions compares favorably to the original Bijaoui's algorithm.

Furthermore, we have developed a methodology to estimate the algorithm parameters (noise power and signal power) allowing for different noise levels at each wavelet scale. This methodology makes it feasible to apply the algorithm in the case of low-pass filtered images resembling

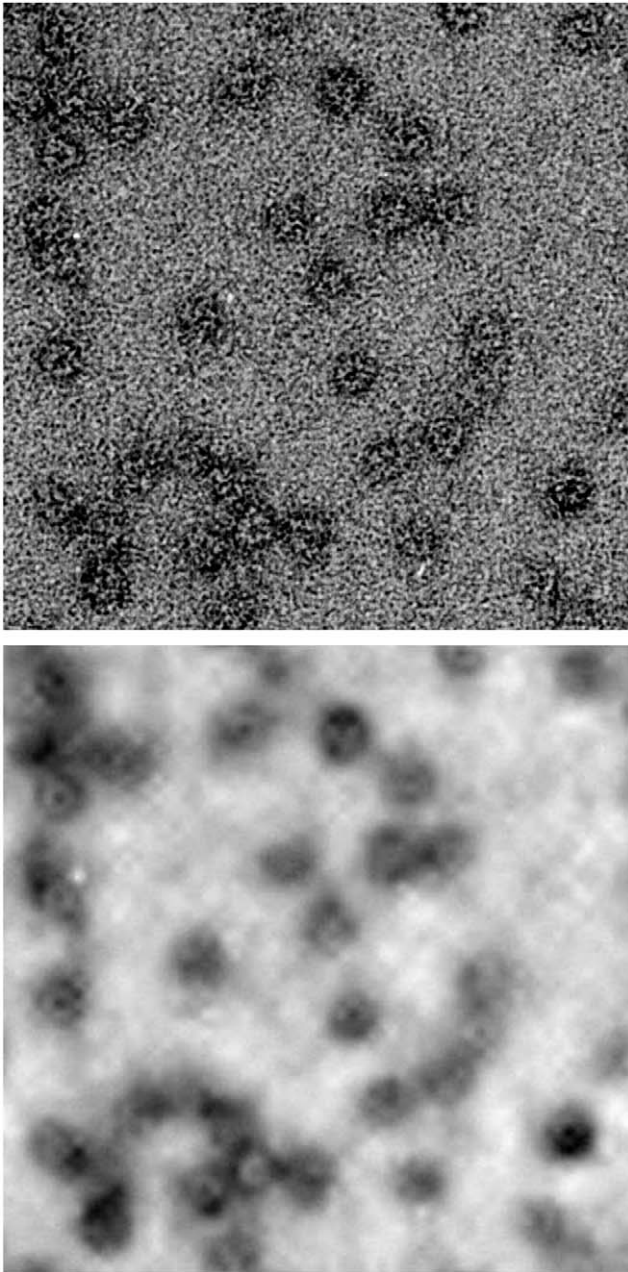


Fig. 6. Top: Negative staining micrograph. Each object in the image correspond to a projection image of the DnaB protein. Bottom: Denoised micrograph using Method 3.

negative-staining conditions in electron microscopy. Our methodology shows a clear advantage in this case with respect to the originally proposed algorithm since the original parameter estimation was not designed to consider the presence of filtered noise. However, the explicit form of the filter does not need to be known in our methodology. We only assume that its amplitude decreases at finer levels of detail which is a plausible hypothesis in a wide range of applications. We have applied our algorithm to the case of experimental negative staining electron micrographs. We have shown that, despite of the complicated nature of the

filter applied by the electron microscope, our methodology is able to enhance the image quality.

The application of our denoising procedure to electron micrographs may facilitate automatic particle identification by improving the image quality as was shown in our experiments. The application of image pre-processing steps like the one described here is important in order to minimize the number of false positives and false negatives [18] selected by the automatic particle-picking algorithms. The selection of thousands of projection images from single particles is an absolute need to achieve high resolution structures. Furthermore, current projects on high-throughput structural proteomics [19] need to solve protein structures in ever decreasing times and need the help of automated steps. The results obtained in this paper encourage further research on this topic.

### Acknowledgements

We acknowledge partial support from the “Comunidad Autónoma de Madrid” through grants CAM-07B-0032-2002 and CAM GR/SAL/0234, the “Comisión Interministerial de Ciencia y Tecnología” of Spain through grants BIO2001-1237, BIO2001-4253-E, BIO2001-4339-E, BIO2002-10855-E, the European Union through grants QLK2-2000-00634, QLRI-2000-31237, QLRT-2000-0136, QLRI-2001-00015, and the NIH through grant HL70472. Partial support from the “Universidad San Pablo-CEU” grant 17/02 is also acknowledged. We are thankful to Dr. Marabini for useful comments on the manuscript.

### References

- [1] A. Bijaoui, Wavelets, Gaussian mixtures and Wiener filtering, *Signal Process.* 82 (2002) 709–712.
- [2] S.G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, New York, 1999.
- [3] J.L. Starck, A. Bijaoui, Filtering and deconvolution by the wavelet transform, *Signal Process.* 35 (1994) 195–211.
- [4] F. Abramovitch, T. Sapatinas, B.W. Silverman, Wavelet thresholding via a Bayesian approach, *J. Roy. Statist. Soc. Ser. B* 60 (1998) 725–749.
- [5] S.G. Chang, B. Yu, M. Vetterli, Spatially adaptive wavelet thresholding with context modeling for image denoising, *IEEE Trans. Image Process.* 9 (2000) 1522–1531.
- [6] E.P. Simoncelli, Bayesian inference in wavelet based models, *Lecture Notes Statist.* 141 (1999) 291–308.
- [7] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Denoising using scale mixtures of Gaussians in the wavelet domain, *IEEE Trans. Image Process.* 12 (2003) 1338–1351.
- [8] M. van Heel, B. Gowen, R. Matadeen, Single-particle electron cryo-microscopy: towards atomic resolution, *Quart. Rev. Biophys.* 33 (2000) 307–369.
- [9] J. Frank, *Three-dimensional Electron Microscopy of Macromolecular Assemblies*, Academic Press, San Diego, CA, 1996.
- [10] J.A. Velázquez-Muriel, C.O.S. Sorzano, J.J. Fernández, J.M. Carazo, A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjusting, *Ultramicroscopy* 96 (2003) 17–35.

- [11] D. Mackay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, Cambridge, 2003.
- [12] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, *Nucleic Acids Res.* 28 (2000) 235–242.
- [13] T.A. Ceska, R. Henderson, J.M. Baldwin, F. Zemlin, E. Beckmann, K. Downing, An atomic model for the structure of bacteriorhodopsin, a seven-helix membrane protein, *Acta Physiol. Scand. Suppl.* 607 (1992) 31–40.
- [14] C.O.S. Sorzano, R. Marabini, J. Velázquez-Muriel, J.R. Bilbao-Castro, S.H.W. Scheres, J.M. Carazo, A. Pascual-Montano, XMIPP: a new generation of an open-source image processing package for electron microscopy, *J. Struct. Biol.* 148 (2004) 194–204.
- [15] W. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C*, second ed., Cambridge University Press, Cambridge, 1992.
- [16] R.E. Walpole, R.H. Meyers, S.L. Meyers, K. Ye, K. Yee, *Probability and Statistics for Engineers and Scientists*, seventh ed., Prentice-Hall, Upper Saddle River, NJ, USA, 2002.
- [17] M. Bárcena, L.E. Donate, T. Ruiz, N. Dixon, M. Radermacher, J.M. Carazo, The DnaB-DnaC complex: a structure based on interactions among asymmetric dimers, *EMBO J.* 20 (2001) 1462–1468.
- [18] W.V. Nicholson, R.M. Glaeser, Review: automatic particle detection in electron microscopy, *J. Struct. Biol.* 133 (2001) 90–101.
- [19] M.B. Schmid, Structural proteomics: the potential of high-throughput structure determination, *Trends Microbiol.* 10 (2001) S27–S31.

**About the author**—C.O.S. SORZANO was born in Málaga (Spain, 1973) where he studied Electrical Engineering (M.Sc., Honors) and Computer Science (B.Sc.). He then joined the BioComputing Unit at the National Center for BioTechnology (CNB) of the Spanish National Council of Scientific Research (CSIC), Madrid, Spain, where he obtained a Ph.D. (Honors). He worked as a Research Assistant in the Biomedical Imaging Group (EPFL, Switzerland), and currently he is a Lecturer in University San Pablo-CEU (Madrid, Spain). His research interests include image processing, tomography, multiresolution approaches, and electron microscopy.

**About the author**—M.D. LÓPEZ was born in Madrid (Spain, 1968) where she studied Mathematics (Astronomy and Geodesy). She has always been related with the University and currently she holds a Lecturer position at the Polytechnical University of Madrid where she joined the research group: Applied Mathematics to the Civil Engineering. Her research interests include: Image Processing, Applied Mathematics to the Engineering, Localization and Computational Geometry.

**About the author**—J. RODRIGO was born in Madrid (Spain, 1967) where he studied Mathematics at the Autonomía University. He has always been related with the University and currently he holds a Lecturer position at the Universidad Pontificia de Comillas (Madrid). He joined the research group: Applied Mathematics to the Civil Engineering. His research interests include: Image Processing, Discrete Mathematics, Localization and Computational Geometry.