

Image restoration methods for the Large Binocular Telescope (LBT)

M. Bertero and P. Boccacci

INFN and DISI, Università di Genova, Via Dodecaneso 35, I-16146 Genova, Italy

Received May 16; accepted September 7, 2000

Abstract. A complete exploitation of the imaging properties of the Large Binocular Telescope (LBT) will require a generalization of the restoration methods which apply to the case of a single image. Several different observations must be combined to obtain a high-resolution representation of a given target. The purpose of this paper is to extend to this problem some of the most used restoration methods, including linear methods such as Tikhonov regularization as well as iterative regularization methods providing positive solution. The proposed methods are implemented and tested on simulated LBT images of diffuse and point-like objects. The results are discussed both from the point of view of the accuracy and from that of the computational efficiency, because LBT images may contain, in principle, up to 10^8 pixels.

Key words: data analysis — methods numerical — image processing — telescopes

1. Introduction

The Large Binocular Telescope (LBT) has been designed for obtaining optical/infrared images with high sensitivity and resolution (Angel et al. 1998). It will consist of two 8.4 m mirrors on a common mount, with a spacing of 14.4 m between the centres. Adaptive optics (AO) will counteract the blurring effect of atmospheric turbulence whereas interferometry will improve the resolution of a single aperture.

For a given orientation of the telescope the diffraction-limited resolution along the centre-to-centre baseline will be equivalent to that of a 22.8 m mirror, while the resolution along the perpendicular direction will be that of a single 8 m mirror. However, a sequence of exposures, taken at sufficiently different parallactic angles, allows in

principle a coverage of the $u - v$ plane equivalent to that of a 22.8 m mirror.

As follows from this remark, an important feature of LBT is that, for a given target, it provides several different interferometric images which must be processed to obtain a high-resolution representation of the target. Thanks to the good coverage of the $u - v$ plane, accurate imaging of complex and diffuse objects is also expected. But these achievements require the solution of a new and interesting problem which is a generalization of the classical restoration (deconvolution) problem and can be called deconvolution of multiple images of the same object.

There exists an extremely wide literature on the classical problem of image deconvolution: one object from one image. A tutorial introduction to this topic can be found in Bertero & Boccacci (1998), hereafter referred to as I. On the other hand the problem of multiple images deconvolution has been considered in a limited number of papers. We quote, for instance, Berenstein & Patrick (1990), Yaroslavsky & Caulfield (1994), Casey & Walnut (1995), Piana & Bertero (1996), as well as Hege et al. (1995), Reinheimer et al. (1997), Correia & Richichi (1999) for specific methods applied to LBT imaging.

The purpose of this paper is twofold: first, to generalize some of the most important deconvolution methods to the case of multiple images, with particular emphasis on LBT imaging; second, to test these methods on a sample of simulated LBT images obtained by assuming perfect AO correction.

In Sect. 2 we formulate the restoration problem of LBT as a least-squares problem and we recall the ill-posedness of this approach. In Sect. 3 we discuss the application of linear regularization methods and, more specifically, of the so-called Tikhonov Regularization (TR) (Engl et al. 1996). In addition we introduce the corresponding *global PSF*, a quantity suitable to describe the outcome of the restoration in terms of resolution. In Sect. 4 we extend to LBT iterative methods such as the Projected Landweber method (PL) (Eicke 1992) and the Image Space Reconstruction Algorithm (ISRA)

Send offprint requests to: M. Bertero,
e-mail: bertero@disi.unige.it

(Daube-Whiterspoon & Muehllhner 1986) which regularize the solution of the least-squares problem with the constraint of positivity. In Sect. 5 we discuss methods derived from Maximum Likelihood in the case of Poisson noise such as the Lucy-Richardson method (LR) (Richardson 1972; Lucy 1974), also known in tomography as Expectation Maximization (EM) (Shepp & Vardi 1982), and an accelerated version of EM, known as Ordered Subsets – Expectation Maximization (OS-EM) (Hudson & Larkin 1995), whose extension to LBT was proposed in Bertero & Boccacci (2000). Finally in Sect. 6 we present the results obtained by applying these methods (TR, PL, ISRA, LR/EM and OS-EM) to a set of simulated LBT images. The accuracy achievable is estimated both in terms of a restoration error (defined as the Euclidean norm of the difference between the restored image and the original object) and in terms of integrated photometry, at least in the case of stellar sources. The computational cost of each method is also estimated. Our conclusions are presented in Sect. 7.

2. Formulation of the problem

If we assume that p observations, corresponding to p different orientations of the LBT baseline, have been performed and that for each observation the PSF is space-invariant, then the mathematical model describing the formation of the p images $g_j(x, y)$, ($j = 1, \dots, p$) is the following

$$g_j(x, y) = \int \int K_j(x - x', y - y') f_0(x', y') dx' dy' + b_j(x, y) + w_j(x, y) \quad (1)$$

where $f_0(x, y)$ is the brightness distribution of the unknown astronomical object, $K_j(x, y)$ is the space-invariant PSF of the LBT interferometer, $b_j(x, y)$ is a function describing the average background due to sky emission and, finally, $w_j(x, y)$ represents the noise (typically Poisson noise and read-out noise) corrupting the j -th image.

The image restoration problem consists in evaluating an estimate $f(x, y)$ of $f_0(x, y)$, being given the detected images $g_j(x, y)$, the PSFs $K_j(x, y)$ as well as estimates of the average backgrounds $b_j(x, y)$. The latter estimates are important in many circumstances. Indeed, if the restoration method includes the positivity constraint, such a constraint is not active without background subtraction. In general one can assume that the functions $b_j(x, y)$ are constant over the field of view of the telescope.

In practice the LBT images are discrete and consist of $N \times N$ pixels, which we characterize by the indices $m, n = 0, 1, \dots, N - 1$. If we denote by $g_j(m, n)$, $f_0(m, n)$ etc. the value of g_j , f_0 etc. at the pixel m, n , then the discrete version of Eq. (1) is given by

$$g_j(m, n) = \sum_{m', n'=0}^{N-1} K_j(m - m', n - n') f_0(m', n') + b_j(m, n) + w_j(m, n). \quad (2)$$

In many circumstances, and precisely when the size of the central peak of the PSF is much smaller than the size of the image, a sufficiently good approximation is obtained if the convolution product of Eq. (2) is intended as a cyclic convolution (periodic PSF), so that, by taking the Discrete Fourier Transform (DFT) of both sides of this equation, we obtain the usual relationship

$$\hat{g}_j(m, n) = \hat{K}_j(m, n) \hat{f}_0(m, n) + \hat{b}_j(m, n) + \hat{w}_j(m, n), \quad (3)$$

where the hat denotes the DFTs. The DFTs of the discrete PSFs $K_j(m, n)$ are the discrete *transfer functions* (TF) of the system corresponding to the different orientations of the baseline.

Since the PSFs of LBT are bandlimited, if they are correctly sampled then the discrete TFs should be zero outside the set of pixels corresponding to the band. However, in practical applications the PSFs of Eq. (2) are measured (for example by imaging a guide star). In such a case the DFTs of these PSFs are, in general, different from zero everywhere, as a consequence of experimental errors and noise. Then the band of the j -th PSF can be defined as the set of the pairs of indices m, n where $\hat{K}_j(m, n)$ is greater than some threshold value related to the experimental errors. We denote this set by \mathcal{B}_j and we assume that $\hat{K}_j(m, n)$ is set to zero outside \mathcal{B}_j .

In this section we formulate the image restoration problem for LBT as a *least-squares problem*. This is equivalent to a Maximum Likelihood approach in the case of white Gaussian noise.

In order to formulate this problem and the related equations in a concise form, we introduce a vector-matrix notation and we write Eq. (2) as follows

$$\mathbf{g}_j = \mathbf{A}_j \mathbf{f}_0 + \mathbf{b}_j + \mathbf{w}_j \quad (4)$$

where \mathbf{g}_j , \mathbf{f}_0 etc. are the arrays formed by $g_j(m, n)$, $f_0(m, n)$ etc. and \mathbf{A}_j is the block-circulant matrix defined by the cyclic convolution with the j -th PSF. Then, if we denote by \mathbf{b}_j an estimate of the j -th background contribution (obtained, for instance, by means of some suitable preprocessing of the detected images \mathbf{g}_j), a *least-square estimate* of \mathbf{f}_0 is any array \mathbf{f} which minimizes the *discrepancy functional*:

$$\varepsilon^2(\mathbf{f}) = \sum_{j=1}^p \|\mathbf{A}_j \mathbf{f} + \mathbf{b}_j - \mathbf{g}_j\|^2 \quad (5)$$

where $\|\cdot\|$ denotes the usual Euclidean norm. This equation implies that it is convenient to introduce *subtracted images* $\mathbf{g}_{s,j}$ defined as follows

$$\mathbf{g}_{s,j} = \mathbf{g}_j - \mathbf{b}_j. \quad (6)$$

By means of standard variational procedures (see I), it is easy to prove that to minimize the functional of Eq. (5) is equivalent to solve the following *normal equation*

$$\sum_{j=1}^p \mathbf{A}_j^T \mathbf{A}_j \mathbf{f} = \sum_{j=1}^p \mathbf{A}_j^T \mathbf{g}_{s,j}, \quad (7)$$

where \mathbf{A}_j^T denotes the transposed of the block-circulant matrix \mathbf{A}_j . Then, by taking the DFT of both sides of this equation one easily gets (the * denotes complex conjugation)

$$|\hat{K}(m, n)|^2 \hat{f}(m, n) = \sum_{j=1}^p \hat{K}_j^*(m, n) \hat{g}_{s,j}(m, n) \quad (8)$$

where

$$|\hat{K}(m, n)|^2 = \sum_{j=1}^p |\hat{K}_j(m, n)|^2. \quad (9)$$

We point out that the band of $|\hat{K}(m, n)|$ is the set \mathcal{B} which is the union of the sets \mathcal{B}_j associated with the PSFs corresponding to the different orientations of the baseline. This set represents the coverage of the $u-v$ plane provided by the observations we are considering.

Equation (8) implies that $\hat{f}(m, n)$ is uniquely determined inside but not determined outside the band \mathcal{B} defined above. In other words the solution of the equation is not unique because the images $\mathbf{g}_{s,j}$ do not provide information about the DFT of the object outside \mathcal{B} . In such a case it is usual to introduce the least-squares solution with minimal Euclidean norm, the so-called *generalized solution* (Golub & van Loan 1983), denoted by \mathbf{f}^\dagger , which is obtained by setting $\hat{f}(m, n) = 0$ outside \mathcal{B} . Therefore the DFT of \mathbf{f}^\dagger is given by

$$\hat{f}^\dagger(m, n) = \sum_{j=1}^p \frac{\hat{K}_j^*(m, n)}{|\hat{K}(m, n)|^2} \hat{g}_{s,j}(m, n) \quad (10)$$

when the pair m, n belongs to \mathcal{B} and is zero elsewhere. By taking the inverse DFT of this array, one gets \mathbf{f}^\dagger .

However the generalized solution is not a satisfactory solution of the restoration problem. In the case of a single image it coincides with the solution provided by the so-called *inverse filter* (Frieden 1975) which, as it is well-known, is completely corrupted by noise propagation as a consequence of the ill-posedness of the image restoration problem (see I for a general discussion). The same result holds true also in the case of multiple images (Piana & Bertero 1996), hence the need of introducing the so-called *regularization methods* (see again I for an introduction).

3. Linear regularization methods and the global PSF

A classical approach to the solution of linear ill-posed problems is the so-called *regularization theory*. In its general form it consists in the minimization of a functional which is obtained from the discrepancy functional by adding a penalty term. This is, in general, the square of the norm of a differential (finite difference) operator acting on the image to be restored. In the case of deconvolution problems no one of the various form of regularization provides an extrapolation of the Fourier transform outside the band of the image and all provide in practice very similar result. Therefore we give here the most simple one which

consists in taking as penalty term the energy of the object and which is usually known as *Tikhonov Regularization* (TR) (Engl et al. 1996). In such a case the functional to be minimized is the following one

$$\Phi_\mu(\mathbf{f}) = \sum_{j=1}^p \|\mathbf{A}_j \mathbf{f} - \mathbf{g}_{s,j}\|^2 + \mu \|\mathbf{f}\|^2 \quad (11)$$

where μ is a positive number called *regularization parameter*. When μ tends to zero this functional tends to the discrepancy functional.

For any given value of μ there exists a unique image $\mathbf{f}^{(\mu)}$ which minimizes the functional and is the unique solution of the normal equation

$$\left(\sum_{j=1}^p \mathbf{A}_j^T \mathbf{A}_j + \mu \mathbf{I} \right) \mathbf{f}^{(\mu)} = \sum_{j=1}^p \mathbf{A}_j^T \mathbf{g}_{s,j}, \quad (12)$$

which can also be easily solved by means of the DFT. The result is

$$\hat{f}^{(\mu)}(m, n) = \sum_{j=1}^p \frac{\hat{K}_j^*(m, n)}{|\hat{K}(m, n)|^2 + \mu} \hat{g}_{s,j}(m, n) \quad (13)$$

with $|\hat{K}(m, n)|^2$ defined in Eq. (9).

This equation defines the family of the regularized solutions. As concerns their dependence on the regularization parameter, we recall that there exists an optimum value of μ , in the sense that it minimizes the restoration error. Several methods have been proposed for its estimation (see I). However, in the practice of LBT imaging, the problem of the choice of the optimal value of μ can be satisfactorily solved in the following way. As follows from Eq. (13), if the DFTs of the images and of the PSFs have been computed and stored, then, for each value of μ , the computation of $\mathbf{f}^{(\mu)}$ requires essentially the computation of one inverse FFT. Therefore the method can be used routinely for a fast inversion of the data and the choice of the best value of μ can be performed interactively by the user.

It is possible to estimate the resolution achievable with a given set of observations without performing explicit data inversion. Such a result can be obtained by generalizing to the LBT problem the concept of *global PSF* introduced in I. The global PSF is the PSF of two linear systems in cascade: the first is the imaging system (the LBT in our case) while the second is the linear filter corresponding to the TR restoration method.

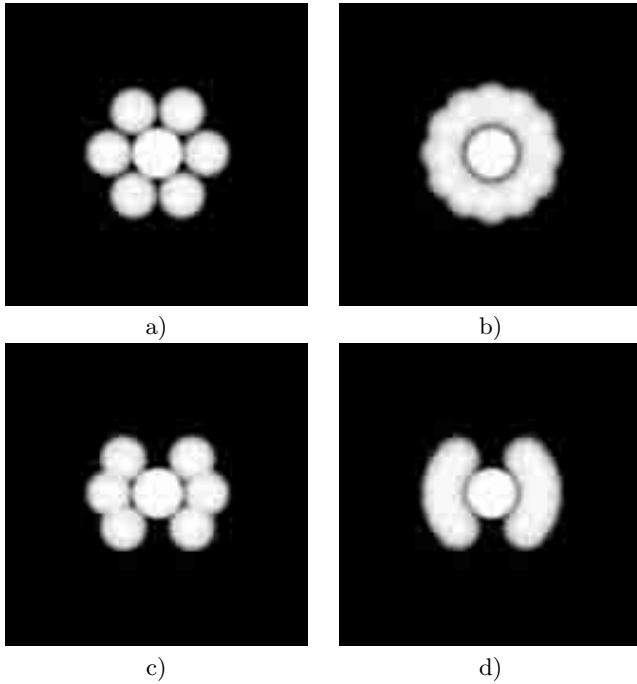


Fig. 1. Gray level representation of the global transfer function of Eq. (15) for various sets of baseline orientations: **a)** three directions spaced by 60° ; **b)** six directions spaced by 30° ; **c)** three directions spaced by 45° ; **d)** six directions spaced by 18° . In the cases **a)** and **c)**, $\mu = 3 \cdot 10^{-3}$ while, in the cases **b)** and **d)**, $\mu = 6 \cdot 10^{-3}$

If we insert the images $\hat{g}_{s,j}(m, n)$, as given by Eqs. (6) and (3), into Eq. (13), we obtain

$$\hat{f}^{(\mu)}(m, n) = \frac{|\hat{K}(m, n)|^2}{|\hat{K}(m, n)|^2 + \mu} \hat{f}_0(m, n) + \sum_{j=1}^p \frac{\hat{K}_j^*(m, n)}{|\hat{K}(m, n)|^2 + \mu} \hat{w}_j(m, n). \quad (14)$$

This equation has a structure similar to that of Eq. (3) after background subtraction. Indeed, the DFT of the restored image $\hat{f}^{(\mu)}$ consists of two terms: the first is the DFT of the object multiplied by the transfer function

$$\hat{K}^{(\mu)}(m, n) = \frac{|\hat{K}(m, n)|^2}{|\hat{K}(m, n)|^2 + \mu}; \quad (15)$$

the second is a contribution due to the noise. Therefore it is quite natural to define $\hat{K}^{(\mu)}(m, n)$ as the *global transfer function* and, accordingly, to define its inverse DFT, $\mathbf{K}^{(\mu)}$, as the *global PSF*.

It is evident that the global transfer function has a support coinciding with the coverage of the $u - v$ plane provided by the observations we use and that, over this support, modulates the Fourier transform of the object in a way depending on the value of μ . We recall that the optimal value of this parameter depends on the amount of noise: smaller noise corresponds to smaller values of μ , hence to more detailed restorations.

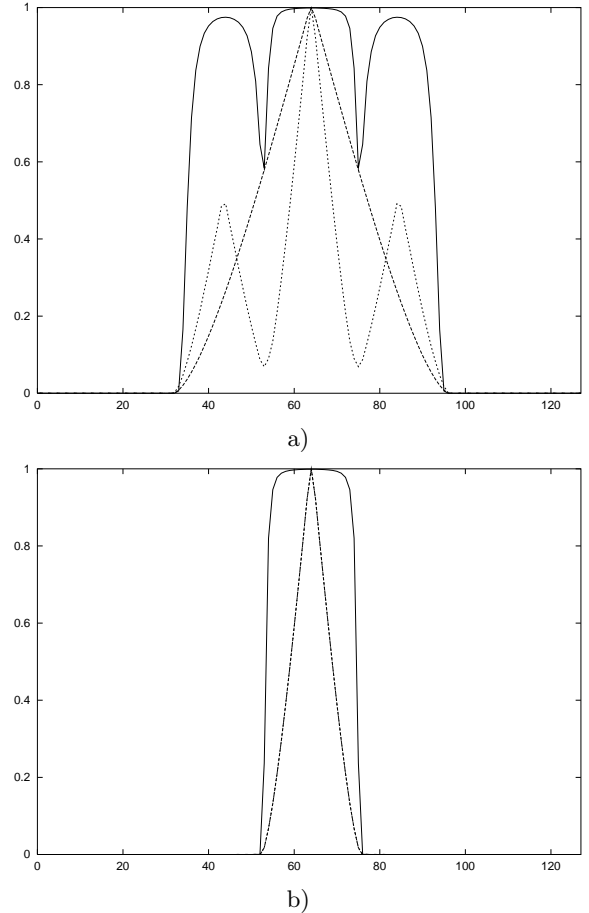


Fig. 2. **a)** Horizontal cut (full line) of the global transfer function of Fig. 1d or Fig. 1b, compared with a cut of the transfer function of a 22.8 m telescope (dashed line) and the cut, along the baseline, of the transfer function of the LBT interferometer (dotted line). **b)** Vertical cut (full line) of the global transfer function of Fig. 1d, compared with the cut, along the direction orthogonal to the baseline, of the transfer function of the LBT interferometer (dash-dotted line). The latter coincides with the cut of the transfer function of a 8.4 m telescope

In order to quantify the previous remarks we consider a few possible situations, by assuming diffraction limited PSFs of the LBT interferometer, i.e. cosine-modulated Airy functions. They are normalized in such a way that their zero frequency component is 1. In addition their FWHM is about 4 pixels in the direction of the baseline and about 11.2 pixels in the orthogonal direction (with the same units the FWHM of a 22.8 m pupil is 5.5 pixels). The interferometric PSFs are computed for p equispaced orientations: in one case they cover 180° while in the other they cover 90° . In both cases we consider $p = 3$ and $p = 6$. As concerns the value of the regularization parameter, from a criterion proposed in Miller (1970) it follows that it is roughly proportional to p . We take $\mu = p \cdot 10^{-3}$, a value which roughly corresponds to the optimal value used for the first numerical example of Sect. 6.

In Fig. 1 we give pictures of the global transfer functions corresponding to the situations indicated above. In these figures the circular bands of the 8.4 m and of the 22.8 m mirror are evident. The lower panels of Fig. 1 show a striking similarity with the situation encountered in limited angle tomography which, as it is well known, does not allow satisfactory image restorations. However the results of Sect. 6 indicate that the situation is more favourable in the case of LBT, mainly as a consequence of the information carried out by the 8.4 m mirror.

In Fig. 2 we plot horizontal and vertical cuts of the global transfer function of Fig. 1d, both through the centre of the inner disc. The horizontal cut, represented in Fig. 2a (full line), is compared with a cut of the transfer function of a 22.8 m telescope (dashed line) and with the cut, along the direction of the baseline, of the transfer function of the LBT interferometer (dotted line). Thanks to the use of a restoration method for producing the final image, the global transfer function of LBT assures a more accurate transmission of the Fourier components of the object than the transfer function of a 22.8 m telescope. As a consequence, in the case of a complete coverage as in Fig. 2b, the FWHM of the global PSF is about 2.8 pixels, against the 5.5 pixels of the PSF of a 22.8 m telescope. The price to be payed is that the central peak of the global PSF is surrounded by rather important negative rings. We observe that similar results could be obtained by applying TR to the images of a hypothetical 22.8 m telescope.

In Fig. 2b the vertical cut of the global transfer function of Fig. 1d (full line) is compared with the cut of the transfer function of the LBT interferometer along the direction orthogonal to the baseline (dash-dotted line). The latter coincides with a cut of the transfer function of a 8.4 m telescope. Again the use of restoration methods implies a more accurate transmission of the Fourier components of the object. In addition, if we compute the global PSF corresponding to the transfer function of Fig. 1d, we find that its FWHM in the vertical direction is about 4 pixels which is much smaller than that of a 8.4 m telescope (11 pixels). This effect may be due to the contributions coming from other directions in the $u-v$ plane and is in agreement with the result of a simulation reported in Sect. 6 (see Fig. 3d).

4. Iterative regularization methods

An important topic in regularization theory is the design of iterative methods with the semiconvergence property, i.e. the iterates first approach the true object \mathbf{f}_0 and then go away toward the generalized solution \mathbf{f}^\dagger . It follows that there exists an optimum number of iterations which can be estimated by extending some of the criteria used for the choice of the regularization parameter.

In the case of the inversion of large data sets these methods are important from the computational point of

view because, at each step, the basic operation is a matrix-vector multiplication. The prototype of these methods is the so-called *Landweber method* (see I), which can also be extended to our problem. It is a linear restoration method (it can be shown that it is equivalent to a linear filtering of the least-squares solutions) and therefore its output can also be described by a global PSF. Its main drawback is the slow convergence and, for this reason, it is not frequently used in practice. Indeed, the most popular iterative method is a nonlinear one, which is called *Conjugate Gradients* (CG) (Golub & van Loan 1983) and provides in practice the same results as Landweber (see I) but with a much smaller number of iterations. CG can also be extended to the problem we are considering but we do not give the algorithm because we do not use it in this paper

One of the shortcomings of the linear methods as well as of CG is that, in the case of sharp intensity variations, they may produce ringing effects. If the background has been subtracted, then the ringing may generate negative values in the restored image. In such a case it is well known that methods generating positive solutions, i.e. methods implementing the so-called *positivity constraint*, may produce a considerable improvement of the restoration. In other words background subtraction may be required to make the positivity constraint to be active.

A quite natural approach is to consider the following least-squares problem

$$\sum_{j=1}^p \|\mathbf{A}_j \mathbf{f} - \mathbf{g}_{s,j}\|^2 = \text{minimum}, \quad \mathbf{f} \geq 0 \quad (16)$$

where $\mathbf{f} \geq 0$ means $f(m, n) \geq 0$ for any m, n . It is easy to show that this problem has always a solution (in general, more than one) and that the set of the solutions is closed and convex, so that there exists a unique solution of minimal Euclidean norm. We call it the *nonnegative generalized solution* and we denote it by \mathbf{f}_+^\dagger . Of course this solution may be corrupted by uncontrolled noise propagation as \mathbf{f}^\dagger is. For this reason the problem is not to compute \mathbf{f}_+^\dagger but a regularized approximation of \mathbf{f}_+^\dagger . To this purpose we can use a modified version of the Landweber method, called the *projected Landweber method* (PL) (Eicke 1992).

For LBT imaging the algorithm is the following

$$\mathbf{f}_+^{(k+1)} = P_+ \left\{ \mathbf{f}_+^{(k)} + \tau \sum_{j=1}^p \left(\mathbf{A}_j^T \mathbf{g}_{s,j} - \mathbf{A}_j^T \mathbf{A}_j \mathbf{f}_+^{(k)} \right) \right\} \quad (17)$$

where k is the number of iterations, τ is a *relaxation parameter* which satisfies the conditions

$$0 < \tau < \frac{2}{\hat{K}_{\max}^2}, \quad (18)$$

\hat{K}_{\max}^2 being the maximum value of the transfer function defined in Eq. (9), and P_+ is the projection operator onto the set of nonnegative arrays, which is defined by: $(P_+ \mathbf{f})(m, n) = f(m, n)$, if $f(m, n) > 0$, $(P_+ \mathbf{f})(m, n) = 0$ if $f(m, n) \leq 0$.

General results on the projected Landweber method (Eicke 1992) assure that, for any initial guess $\mathbf{f}_+^{(0)}$, the iterates $\mathbf{f}_+^{(k)}$ converge to a solution of the constrained least-squares problem defined in Eq. (16). The usual choice is $\mathbf{f}_+^{(0)} = 0$ because, in such a case, it is expected to obtain a regularized approximation of \mathbf{f}_+^\dagger by means of a suitable stopping of the iterations. Indeed, numerical experience indicates that the method has the semiconvergence property. The main difficulty is that the optimum number of iterations is, in general, very large. On the other hand each iteration is very fast because it only requires the computation of two FFTs, one direct and one inverse.

Another iterative method for the solution of the constrained least-squares problem of Eq. (16) is the so-called Iterative Space Reconstruction Algorithm (ISRA), proposed by Daube-Whiterspoon & Muehllhner (1986). The proof of the convergence of the algorithm is given in De Pierro (1987). It can be formulated in a concise form by defining the product of two arrays \mathbf{g} and \mathbf{h} , denoted by \mathbf{gh} , as the array given by $(\mathbf{gh})(m, n) = g(m, n)h(m, n)$ and their quotient $\mathbf{g/h}$ as the array given by $(\mathbf{g/h})(m, n) = g(m, n)/h(m, n)$ (provided that $h(m, n)$ is different from zero everywhere). With these notations the extension of ISRA to the LBT problem is the following

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} \frac{\sum_{j=1}^p \mathbf{A}_j^T \mathbf{g}_{s,j}}{\sum_{j=1}^p \mathbf{A}_j^T \mathbf{A}_j \mathbf{f}^{(k)}}. \quad (19)$$

It must be observed that the method applies to the case of nonnegative images and nonnegative PSFs, so that, if we start with a positive $\mathbf{f}^{(0)}$ (for instance a uniform array), then all iterates are nonnegative. But, in our applications, the subtracted images $\mathbf{g}_{s,j}$ may take negative values after background subtraction. These negative values may generate negative values in $\mathbf{A}_j^T \mathbf{g}_s$ and consequently instabilities of the algorithm. For this reason, the negative values of $\mathbf{A}_j^T \mathbf{g}_{s,j}$ must be set to zero and these zeros may propagate to all $\mathbf{f}^{(k)}$ s.

As concerns the regularization effect of the algorithm, our simulations indicate that it has the semiconvergence property and that the optimal number of iterations is in general very large also in this case. The computational cost per iteration is the same of PL. The advantage of PL is that it is more flexible and can be used for introducing other constraint besides positivity (see I).

5. Maximum Likelihood methods in the case of Poisson noise

The least-squares methods discussed in the previous sections derive from the assumption that the noise is white and Gaussian, even if they work well also when this assumption is not satisfied. In the case of Poisson noise Maximum Likelihood provides one of the most popular methods for the restoration of astronomical images,

the so-called Lucy-Richardson method, which is a particular version of the more general Expectation Maximization method. We denote as LR/EM its extension to LBT imaging, which is given in Correia & Richichi (1999) and Bertero & Boccacci (2000). This extension however, applies to the case where the background is zero or negligible. If this condition is not satisfied, it is not possible to use the subtracted images defined in Eq. (6) because they may contain a large number of pixels with negative values. Zeroing of these values may change significantly the total number of counts.

The correct approach is obtained by modeling the images \mathbf{g}_j of an astronomical object \mathbf{f} as translated Poisson processes (Snyder & Miller 1991) with expectations given by $\mathbf{A}_j \mathbf{f} + \mathbf{b}_j$. Here \mathbf{b}_j is the expectation of the independent Poisson process describing the background of the j -th observation. If all the observations are statistically independent, then the log-likelihood function is given by

$$l(\mathbf{f}) = \sum_{j=1}^p \sum_{m,n=0}^{N-1} \{g_j(m, n) \ln[(\mathbf{A}_j \mathbf{f})(m, n) + b_j(m, n)] - (\mathbf{A}_j \mathbf{f})(m, n)\}. \quad (20)$$

Using the notations introduced in Sect. 5 and repeating standard arguments for the derivation of the EM algorithm (Shepp & Vardi 1982), we can find that LR/EM is given by

$$\eta \mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} \sum_{j=1}^p \mathbf{A}_j^T \frac{\mathbf{g}_j}{\mathbf{A}_j \mathbf{f}^{(k)} + \mathbf{b}_j} \quad (21)$$

where

$$\eta = \sum_{j=1}^p \hat{K}_j(0, 0). \quad (22)$$

As in the case of ISRA the algorithm is usually initialized with a uniform image $\mathbf{f}^{(0)}$.

The algorithm defined by Eq. (21) provides a sequence converging to a maximum point of the log-likelihood function. However it is well-known that it is not convenient to push the algorithm too far because too many iterations produce noise amplification. This property is similar to that of the iterative methods for the least-squares problem. Numerical simulations indicate that also the LR/EM method has the semiconvergence property. In the case of real images criteria for stopping the iterations are required (Lantéri et al. 1999). Finally, as concerns the implementation of the method, one iteration requires the computation of $3p+1$ FFTs so that the computational cost of one iteration is much higher than that of one PL or ISRA iteration.

An annoying feature of the LR/EM method is its slow convergence. This property was remarked also in the case of computed tomography where an accelerated version of EM, known as Ordered Subsets-EM (OS-EM) was proposed by Hudson & Larkin (1995). This accelerated version has been adapted by the authors (Bertero & Boccacci 2000) to the restoration of LBT images. We give here the

version corresponding to the algorithm of Eq. (21), i.e. the version which takes properly into account the background contribution.

The proposed version of the OS-EM method is as follows:

- initialize with $\mathbf{f}^{(0)}$ positive (for instance a uniform image);
- given $\mathbf{f}^{(k)}$ set $\mathbf{h}^{(0)} = \mathbf{f}^{(k)}$ and, for $j = 1, 2, \dots, p$, compute

$$\eta_j \mathbf{h}^{(j)} = \mathbf{h}^{(j-1)} \mathbf{A}_j^T \frac{\mathbf{g}_j}{\mathbf{A}_j \mathbf{h}^{(j-1)} + \mathbf{b}_j}; \quad (23)$$

- set $\mathbf{f}^{(k+1)} = \mathbf{h}^{(p)}$.

In Bertero & Boccacci (2000) we demonstrate, by means of numerical simulations, that this method provides in practice the same results as LR/EM, with a reduction in the number of iterations by a factor p . Since the computational cost of one OS-EM iteration (computation of $4p$ FFTs) is not much higher than that of one LR/EM iteration (computation of $3p + 1$ FFTs), it follows that OS-EM assures a reduction of the computational cost by a factor of about p . It also follows that the total computation time for a restoration by means of OS-EM is approximately independent of the number of images p . Hence OS-EM may be very convenient for the deconvolution of large images as those which will be provided by LBT, even if the faster convergence implies that the choice of the stopping rule is more critical than in the case of LR/EM. This point will deserve an accurate investigation.

6. Numerical results

We have implemented the five restoration methods discussed in the previous sections: TR, as the prototype of the linear methods used for regularizing least-squares solutions; PL and ISRA, as regularization methods for constrained least-squares solutions with the positivity constraint; LR/EM and its accelerated version OS-EM. The methods TR, PL and ISRA apply, in principle, to the case of white Gaussian noise but are commonly used also for the restoration of images corrupted by Poisson and Gaussian noise.

All the algorithms include one or more parameters whose values must be optimized by means of numerical simulations, according to the tasks established by the users. This procedure may be called the *training* of the algorithm and the tasks may be defined through the so-called *Figures of Merit* (FOM).

In this paper, we use three different FOMs. The first reflects the ability of the algorithms to minimize the *restoration error* (RE), i.e. the relative RMS discrepancy between the true object \mathbf{f}_0 and its restored version \mathbf{f}_{res}

$$RE = \frac{\|\mathbf{f}_{\text{res}} - \mathbf{f}_0\|}{\|\mathbf{f}_0\|}. \quad (24)$$

In addition we introduce two different photometric FOMs to evaluate the photometric efficiency of the various methods. The first is related to the flux on a set of *Regions Of Interest* (ROI) while the second evaluates the photometric accuracy in the case of a single star or of a cluster of stars.

If the ROIs are R_1, R_2, \dots, R_r , then in each one the fluxes of the original image and of the restored one are given by:

$$\Phi_l = \sum_{m,n \in R_l} f_0(m,n); \quad \Phi_l^{(\text{res})} = \sum_{m,n \in R_l} f_{\text{res}}(m,n) \quad (25)$$

with $l = 1, 2, \dots, r$, so that we can define the *Flux Error* (FE) as follows:

$$FE = \frac{1}{r} \sum_{l=1}^r \frac{|\Phi_l - \Phi_l^{(\text{res})}|}{\Phi_l}. \quad (26)$$

As concerns the second photometric FOM, let us assume that we are considering a set of s stars corresponding to the regions S_1, S_2, \dots, S_s . Then, for each star, we can compute the magnitude as follows:

$$m_l = -2.5 \log_{10} \left(\sum_{m,n \in S_l} \frac{f_0(m,n)}{f_b} \right) \quad (27)$$

($l = 1, 2, \dots, s$), where f_b is the sky flux per (arcsec)². If we compute the magnitude of the reconstructed star $m_l^{(\text{res})}$ in a similar way, we define the *Average Magnitude Error* (AME) as follows

$$AME = \frac{1}{s} \sum_{l=1}^s |m_s - m_s^{(\text{res})}|. \quad (28)$$

In our numerical simulations we use the same objects and images of previous papers (Correia & Richichi 1999; Bertero & Boccacci 2000). Therefore perfect AO correction is assumed and the interferometric PSFs are cosine-modulated Airy functions (whose FWHM in the direction of the baseline is about 4 pixels). The simulated observations, performed in R band, are obtained by convolving the diffraction-limited PSFs with the object and by adding suitable background and noise (Poisson plus read-out) to the result. Details on the values of the relevant parameters (integration times, sky brightness etc.) can be found in Correia & Richichi (1999).

The first object is an image 128×128 of the spiral galaxy NGC 1288, which is shown in Fig. 3a, where the white squares indicate the ROIs used for the computation of FE . For this object we have computed three sets of simulated images corresponding to 4, 6 and 8 equispaced values of the parallactic angle between 0° and 180° . Since the magnitude of the galaxy is set to $m_r = 19$, this leads to approximately $2 \cdot 10^7$ photons per long exposure images and a peak SNR of 80.

For this example we have investigated the behaviour of two FOMs, RE and FE . As functions of the regularization parameter (TR) or of the number of iterations (PL, ISRA, LR/EM, OS-EM) both exhibit a minimum (hence an optimal restoration from the point of view of that FOM).

Table 1. Summary of the results obtained for the galaxy NGC 1288 by means of the methods implemented in this work. For each method we report the minimum value of RE and FE and the corresponding value of the regularization parameter μ_{opt} or of the number of iterations k_{opt}

method	p	k_{opt} or μ_{opt}	RE	k_{opt} or μ_{opt}	FE
TR	4	$6.5 \cdot 10^{-3}$	0.049	$8.0 \cdot 10^{-4}$	0.0160
	6	$7.3 \cdot 10^{-3}$	0.049	$3.5 \cdot 10^{-3}$	0.0120
	8	$8.8 \cdot 10^{-3}$	0.043	$3.7 \cdot 10^{-3}$	0.0040
PL	4	278	0.048	273	0.0137
	6	329	0.044	361	0.0139
	8	392	0.042	467	0.0024
ISRA	4	373	0.046	408	0.0157
	6	487	0.042	569	0.0126
	8	519	0.040	670	0.0030
LR/EM	4	314	0.044	295	0.0138
	6	380	0.040	381	0.0126
	8	393	0.039	411	0.0059
OS-EM	4	79	0.044	76	0.0140
	6	64	0.040	62	0.0131
	8	50	0.039	52	0.0050

The minima of RE and FE do not occur for the same value of the parameters. The results are summarized in Table 1.

A few comments on these results are appropriate. First, the restoration error is rather small (as a consequence of the good SNR of the simulated images as well as of the good coverage of the $u-v$ plane) and is approximately the same for all methods. Second, the restoration error decreases for increasing number of images, as it is expected, even if the improvement is not very spectacular (a gain of 0.5% from 4 to 8 images). We also notice that the optimal value of the regularization parameter, as well as the optimal value of the number of iterations, increases for increasing number of images (except for OS-EM). An argument justifying this behaviour is given in Sect. 3. In the case of OS-EM, as already shown in Bertero & Boccacci (2000) the number of iterations decreases when the number of images increases, in such a way that the computation time does not significantly depends on the number of images.

Similar remarks apply also to the case of FE , which measures the local accuracy of the restoration. FE is always smaller than RE and, in general, this result is obtained with a smaller value of the regularization parameter or a higher number of iterations. In addition we observe a strong reduction of FE when using eight images instead of four or six.

As already remarked all methods provide in practice the same restoration error. This result is due to the fact that we are restoring a diffuse object with a rather large (and varying) background, so that the positivity

Table 2. Restoration errors corresponding to different values of the total parallactic angle. The results are obtained for the galaxy NGC 1288, using four equispaced images for each value of the total angle

Total angle	(TR) μ_{opt}	(TR) RE	(PL) k_{opt}	(PL) RE	(LR/EM) k_{opt}	(LR/EM) RE
135°	$4 \cdot 10^{-3}$	0.049	278	0.044	314	0.044
93°	$7 \cdot 10^{-3}$	0.053	256	0.052	288	0.048
63°	$7 \cdot 10^{-3}$	0.056	253	0.056	309	0.051

constraint is not active (see I for a discussion). In similar cases one can choose the restoration method on the basis of the computational cost.

We have used our codes also for testing, on the particular example of the galaxy NGC 1288, the effect of incomplete coverage of the $u-v$ plane. We have assumed four equispaced images within total parallactic angles (difference between the angles of the two extreme orientations of the baseline) of 135° , 93° and 63° (the first case corresponds to the first case of Table 1). In Table 2 we report, for simplicity, only the results obtained with TR, PL and LR/EM. The notations are those used in Table 1. Again all methods provide essentially the same restoration error and it is rather surprising to find that this error does not increase dramatically when the total parallactic angle decreases. We remark that the case of 93° corresponds approximately to the coverage of the $u-v$ plane shown in Fig. 1d (after a rotation of about 45°).

A visual confirmation of the result is provided by Fig. 3 where two examples of restorations obtained by means of LR/EM are shown. Figure 3c is obtained using 8 equispaced images covering a total parallactic angle = 180° while Fig. 3d is obtained with 4 equispaced images covering a total parallactic angle = 63° . These pictures suggest that an incomplete coverage of the $u-v$ plane may not influence dramatically the general quality of the restored image, even if details recovered in the case of complete coverage are not recovered in the other one.

The second example we consider is a simulation of binary stars of different relative magnitude. In the synthetic object each star is located in one pixel, with a separation of about 14 pixels which is about 3.5 times the diffraction limit in the direction of the baseline. Moreover two cases are investigated: 1) the magnitudes of the two stars are 27.5 and 30 and the average peak SNR in the images is 11.3 for the main star and 3.5 for the companion; 2) the magnitudes are 29 and 30 while the corresponding average peak SNRs are 5.5 and 3.5 respectively. Also for these examples we have computed sets consisting of 4, 6 and 8 equispaced observations. Moreover, for TR, PL and ISRA the constant background has been subtracted from the simulated LBT images to obtain the subtracted images

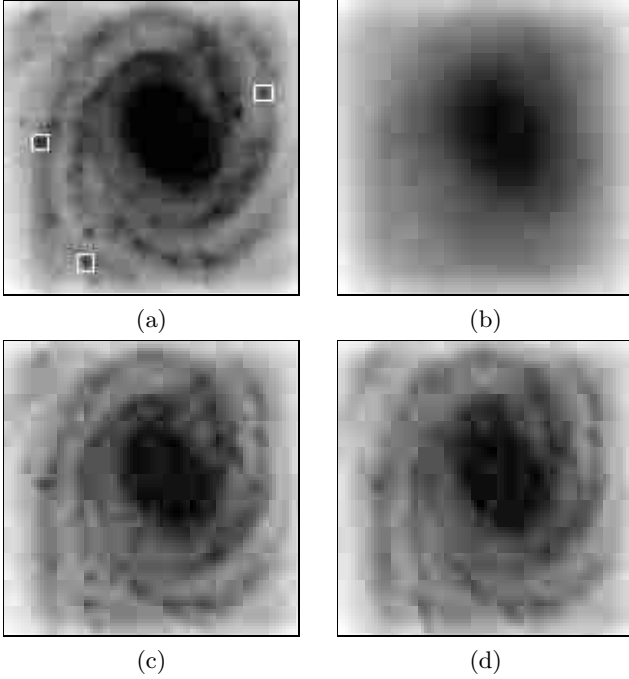


Fig. 3. **a)** Original image of the galaxy NGC 1288. The white squares indicate the ROIs used to evaluate the FOM defined in Eq. (54). **b)** One of the simulated LBT images (parallaxic angle = 0°). **c)** Best restoration provided by the LR/EM method, using 8 equispaced images covering an angle = 180° . **d)** Best restoration provided by the LR/EM method, using 4 images covering an angle = 63°

defined in Eq. (6). In the case of ISRA the negative values of $\mathbf{A}_j^T \mathbf{g}_{s,j}$ have been set to zero. As concerns LR/EM and OS-EM, background subtraction is not needed because the value of the background is inserted directly in the algorithm, as indicated in Sect. 5. However, since the background is not very large, the computed images \mathbf{g}_j can take small negative values in a few pixels as a consequence of the white Gaussian noise simulating the read-out noise. These values are set to zero in order to avoid instabilities of the algorithm.

The five implemented methods have been applied to the three sets of observations for each one of the two cases. The FOMs considered are RE and AME . The results are reported in Table 3 for the first case and in Table 4 for the second one. From the point of view of RE , the number of iterations needed for reaching the minimum is much larger than in the case of the galaxy NGC 1288. This is a general feature of iterative methods: point-like objects need much more iterations than diffuse objects. In addition the minima are usually very flat and, for this reason, in many cases we stopped the procedure after 1000 iterations. The results we have obtained may be summarized by saying that only in the case of LR/EM and OS-EM the restoration error RE decreases when the number of images increases. These two methods also provide a much

Table 3. Summary of the results obtained for the binary star 27.5 – 30 by means of the methods implemented in this work. The * indicates that the quoted number of iterations does not correspond to the minimum of the FOM (see the text)

method	p	k_{opt} or μ_{opt}	RE	k_{opt} or μ_{opt}	AME
TR	4	$1.0 \cdot 10^{-3}$	0.922	$1.2 \cdot 10^{-2}$	0.0294
	6	$1.0 \cdot 10^{-3}$	0.913	$1.6 \cdot 10^{-2}$	0.0375
	8	$1.0 \cdot 10^{-3}$	0.911	$2.5 \cdot 10^{-2}$	0.0345
PL	4	1000*	0.753	1000*	0.1080
	6	1000*	0.747	223	0.0579
	8	1000*	0.753	234	0.0362
ISRA	4	1000*	0.759	516	0.0501
	6	1000*	0.897	520	0.0109
	8	1000*	0.674	931	0.0072
LR/EM	4	1000*	0.063	254	0.0448
	6	1000*	0.047	223	0.0064
	8	1000*	0.032	158	0.0008
OS-EM	4	1000*	0.043	64	0.0439
	6	1000*	0.035	32	0.0099
	8	376	0.024	16	0.0029

Table 4. Summary of the results obtained for the binary star 29 – 30 by means of the methods implemented in this work. The * indicates that the quoted number of iterations does not correspond to the minimum of the FOM (see the text)

method	p	k_{opt} or μ_{opt}	RE	k_{opt} or μ_{opt}	AME
TR	4	$1.4 \cdot 10^{-2}$	0.956	$1.8 \cdot 10^{-2}$	0.0198
	6	$1.2 \cdot 10^{-2}$	0.934	$2.8 \cdot 10^{-2}$	0.0054
	8	$1.3 \cdot 10^{-2}$	0.936	$2.6 \cdot 10^{-2}$	0.0128
PL	4	1000*	0.769	199	0.0412
	6	1000*	0.756	162	0.0441
	8	1000*	0.765	221	0.0166
ISRA	4	1000*	0.394	135	0.0429
	6	1000*	0.231	82	0.0380
	8	1000*	0.334	97	0.0694
LR/EM	4	1000*	0.191	268	0.0310
	6	1000*	0.193	266	0.0323
	8	1000*	0.152	186	0.0350
OS-EM	4	1000*	0.148	69	0.0187
	6	1000*	0.142	44	0.0300
	8	1000*	0.128	23	0.0454

smaller value of RE than the three others. Moreover RE is smaller in the case of example 1) (that with the higher SNR) than in the case of example 2), as it is expected.

As concerns AME , defined in Eq. (28), it has been computed on squares of 5×5 pixels, centred on the pixels of the two stars. The results are shown again in Table 3 and Table 4 for the two cases. It is not always true that the value of AME decreases with increasing number of iterations, even if it is true for LR/EM and OS-EM in the

case of the binary with the higher SNR. In the case of low SNR, quite surprisingly TR is the method providing the best results. In general we observe fluctuations across the various methods but a very promising result is that, for reaching the minimum of AME , we need a quite small number of iterations, especially in the case of OS-EM.

The fluctuations indicated above may be due to the fact that, even if AME is small (the reported values correspond to an error on the fourth significant digit), it may be strongly influenced by noise. Therefore, in order to check this point, we have evaluated the noise dependence of AME by restoring images obtained with five different noise realizations (including that corresponding to the results reported in Table 4) and by computing the minimum value of AME for all these cases.

We restrict the analysis to the binary of Table 4, assuming six observations and using only ISRA, LR/EM and OS-EM. It turns out that the value of AME is always small (in general smaller than the value reported in Table 4) but strongly affected by the change in noise realization. For each method we have computed the average value and standard deviation of AME and the results are the following: ISRA, 0.022 ± 0.012 ; LR/EM, 0.017 ± 0.012 ; OS-EM, 0.015 ± 0.011 . The three methods provide essentially the same results and, in all cases, the standard deviation of AME is approximately equal to its average value.

7. Concluding remarks

In this paper we have investigated the extension to LBT of five image restoration methods. If we take into account both accuracy and efficiency, we can conclude that at least three methods, TR, PL and OS-EM, should be considered as routine method for the analysis of LBT images. We recall that efficiency is an important issue because the wide field of the telescope implies that each image may contain in principle up to 10^8 pixels.

TR is certainly the most efficient method, providing very fast restorations which can also be quite accurate in the case of diffuse objects. Among the iterative methods for positive solutions, PL looks in general more efficient than ISRA, with approximately the same accuracy, while OS-EM is certainly much more efficient than LR/EM. As concerns the comparison between PL and OS-EM the last one provides in some cases better restorations with a much smaller number of iterations. However one must take into account that the ratio between the computational cost of one OS-EM iteration and that of one PL iteration is about $2p$. If we correct the number of iterations by this factor, then PL is a bit more efficient than OS-EM. In addition PL is more flexible and can be modified for taking into account other constraints (for instance the support of localized objects to obtain superresolution effects – see I or Bertero & De Mol 1996). Therefore we think that it must be seriously considered for LBT imaging.

A few words about two popular methods not considered in this paper: Maximum Entropy (ME) and CLEAN. As proved in the beautiful papers of Donoho and coworkers (Donoho et al. 1991), ME provides good results in the case of “nearly-black objects”, namely objects which are zero almost everywhere except in a few pixels. ME can also be easily extended to the case of multiple images. However it is known that the computational cost of ME is high even in the case of a single image so that it is obviously higher in the case of multiple images. Therefore we think that it should be considered for some specific objects as star clusters. As concerns CLEAN, it is well-known that, in the case of a single image, it works well for point-like objects. In the case of multiple images the positions of the maxima in the different images may not coincide and, as a consequence, the extension of CLEAN to this case may be difficult.

Several problems concerning LBT imaging must still be investigated. First, the methods proposed in this paper must be tested on a wider sample of simulated images, using more realistic PSFs corresponding to partial AO correction. Moreover the PSFs are in general space-variant, mainly as a consequence of different AO corrections on the field and therefore deconvolution methods based on a domain decomposition of the images are needed (Fish et al. 1996; Nagy & O’Leary 1998). Finally the PSFs mentioned above are measured by observing one or more guide stars and consequently are corrupted by noise and various kinds of errors so that removal of this degradation may be necessary before their use in image restoration. It follows that the use of (in general nonlinear) denoising methods and/or blind deconvolution methods is required.

The IDL codes of the methods implemented for this paper are free and available on request to the authors. Additional details on these methods are contained in the web page <http://dirac.disi.unige.it>

Acknowledgements. This work was partially funded by CNAA (Consorzio Nazionale per l’Astronomia e l’Astrofisica) under the contract No. 16/97. We thank Maurizio Busso, Mario Gai, Mario Lattanzi, from Osservatorio Astronomico di Torino, Serge Correia, Andrea Richichi and Piero Salinari, from Osservatorio Astrofisico di Arcetri, for several useful discussions and suggestions. We also thank Serge Correia and Andrea Richichi for providing the images used in this paper to test the implemented restoration methods.

References

- Angel J.R.P., Hill J.M., Strittmatter P.A., Salinari P., Weigelt G., 1998, in: *Astronomical Interferometry*, Reasenberg R.D. (ed.), Proc. SPIE 3352, 881
- Berenstein C.A., Patrick E.V., 1990, Proc. IEEE 78, 723
- Bertero M., Boccacci P., 1998, *Introduction to Inverse Problems in Imaging*. IOP Publishing, Bristol
- Bertero M., Boccacci P., 2000, A&AS 144, 181
- Bertero M., De Mol C., 1996, *Progress in Optics XXXVI*, Wolf E. (ed.). Elsevier, Amsterdam

- Casey S.D., Walnut D.F., 1995, *SIAM Rev.* 36, 537
- Correia S., Richichi A., 2000, *A&AS* 141, 301
- Daube-Whiterspoon M.E., Muehlehner G., 1986, *IEEE Trans. Med. Im.* 5, 61
- De Pierro A.R., 1987, *IEEE Trans. Med. Im.* 6, 174
- Donoho D.L., Johnstone I.M., Hock J.C., Stern A.S., 1991, *J. R. Stat. Soc.* 54, 1
- Eicke B., 1992, *Num. Funct. Anal. Opt.* 13, 413
- Engl H.W., Hanke M., Neubauer A., 1996, *Regularization of Inverse Problems*. Kluwer, Dordrecht
- Fish D.A., Grochmalicki J., Pike E.R., 1996, *J. Opt. Soc. Am. A* 13, 464
- Frieden B.R., 1975, in: *Picture Processing and Digital Filtering*, Huang T.S. (ed.), *Topics in Applied Physics*, Vol. 6. Springer, Berlin, p. 179
- Golub G.H., van Loan C.F., 1983, *Matrix Computations*. MD: Johns Hopkins University Press, Baltimore
- Hege E.K., Angel J.R.P., Cheselka M., Lloyd-Hart M., 1995, in: *Advanced Imaging Technologies and Commercial Applications*, Proc. SPIE 2566, Clark N., Gonglewski J.D. (eds.), p. 144
- Hudson H.M., Larkin R.S., 1995, *IEEE Trans. Med. Im.* 13, 601
- Lantéri H., Soummer R., Aime C., 1999, *A&AS* 140, 235
- Lucy L.B., 1974, *AJ* 79, 745
- Miller K., 1970, *SIAM J. Math. Anal.* 1, 52
- Nagy J.G., O'Leary D.P., 1998, *SIAM J. Sci. Comput.* 19, 1063
- Piana M., Bertero M., 1996, *J. Opt. Soc. Am.* 13A, 1516
- Reinheimer T., Hofmann K.-H., Scholler M., Weigelt G., 1997, *A&AS* 121, 191
- Richardson W.H., 1972, *J. Opt. Soc. Am.* 62, 55
- Shepp L.A., Vardi Y., 1982, *IEEE Trans. Med. Imaging MI-1*, 113
- Snyder D.L., Miller M.I., 1991, *Random Point Processes in Space and Time*. Springer Verlag, Berlin
- Yaroslavsky L.P., Caulfield H.J., 1994, *Appl. Opt.* 33, 2157