BAYESIAN DECONVOLUTION IN OPTICAL ASTRONOMY

R. MOLINA
Departamento de Ciencias de la Computación e I.A. Universidad de Granada, 18071 Granada, Spain

ASCENSIÓN DEL OLMO AND JAIME PEREA
Instituto de Astrofísica de Andalucía, Apdo 2144, 18080-Granada, Spain

B. D. RIPLEY
Department of Statistics, University of Oxford, Oxford OX1 3TG, United Kingdom

Received 8 October 1990; revised 9 August 1991

ABSTRACT

In this work we use Bayesian methods and spatial stochastic processes in the deconvolution of images of galaxies. Under very simple but realistic prior assumptions about the true underlying image of a galaxy the Bayesian framework is put to work. The method is tested in CCD images of extragalactic objects of different morphological types and an analysis of the deconvolutions obtained is performed emphasizing the comparison with other observational results.

1. INTRODUCTION

The distortion produced by the Earth’s atmosphere on the observed light distribution of extended objects is the main problem for the spatial resolution in optical astronomy. The observed luminosity distribution of a distant object is the result of the two-dimensional convolution of the unknown light distribution with instrumental effects (telescope + detector) and, mainly, atmospheric distortion. All the processes that produce a blurred image can be represented by a point spread function (PSF). In addition, the collected image is degraded by noise produced by detectors and by the intrinsic discrete nature of light. A clear example of an effect produced by observational distortion is the change in the shape of the brightness profile in cores of elliptical galaxies and in bulges of spirals. Several authors (Schweizer 1979; Frandsen & Thomsen 1980; Nieto 1980; Djorgovsk 1983; and many others) have shown the importance of taking into account the distortion in the determination of physical observational quantities (core radii, sizes, masses, etc.).

Modern detectors provide images of high signal-to-noise ratio and long dynamic range, facilitating the digital processing and the application of restoration methods to remove the seeing blur and enhance small features in the images. During recent years several techniques for the deconvolution of observational PSFs have been developed based on different algorithms, linear or nonlinear; the maximum entropy method (see, e.g., Narayan & Nityanada 1986, and references therein; Jansson et al. 1970; Lucy 1974), the Wiener filter (Arp & Lorr 1976) and other image restorations as the CLEAN algorithm (Högbom 1974) used mainly in radio astronomy.

The quality of a restoration algorithm may be defined in terms of its robustness to the presence of noise and its capability to obtain stable deconvolved images, i.e., it does not restore and does not detail the noise of the observed image (Wells 1980).

In this paper we present a Bayesian method to deconvolve astronomical images in the presence of noise. In Sec. 2 we describe the framework, the Bayesian paradigm, and develop the theory. Section 3 describes and justifies the prior model we will use in the deconvolution of galaxies. Section 4 studies the noise process for our deconvolution problem. Once the prior and the noise processes have been defined, we derive the equation of the maximum a posteriori estimation in Sec. 5. Section 6 is devoted to the study of how robust statistics can be used to solve the problems that arise when we have hot pixels, bad lines, and so on. Section 7 studies how to estimate the only remaining parameter in our model, the variance of the prior model. The choice of the starting point for the iterative method is described in Sec. 8. Computational aspects of the model are discussed in Sec. 9. Finally, in Sec. 10 we study the performance of our method on synthetic and observed images of galaxies.

2. BAYESIAN METHODS

The philosophy within statistics known as Bayesian interference has a very long history, but has only recently been used explicitly in image processing. It is distinguished from the perhaps more familiar classical statistical ideas by using prior information about the images being studied. In our applications this will involve considering the image of being made up of point sources (stars) and smoothly varying luminosity (galaxies). Our methods are distinguished from others most particularly by our inclusion of specific spatial smoothness information about the objects present in the image.

We will distinguish between S, the “true” image which would be observed under ideal conditions (i.e., no noise and no distortions produced by atmospheric blurring and instrumental effects), and Z, the observed image. The aim is then to reconstruct S from Z. Bayesian methods start with a prior distribution, a probability distribution over images S. (It is here that we incorporate information on the expected structure within an image.) It is also necessary to specify P(Z|S), the probability distribution of observed images Z if S were the “true” image. The Bayesian paradigm dictates that interference about the true S should be based on P(S|Z) given by

$$P(S|Z) = P(Z|S)P(S)/P(Z) \propto P(Z|S)P(S). \quad (1)$$

To show just one restoration, it is common (but not obligatory) to choose the mode of P(S|Z), that is, to display the image S which satisfies

---

1 Visiting Astronomers, German–Spanish Astronomical Center, Calar Alto, operated by the Max-Planck-Institut für Astronomie Heidelberg jointly with the Spanish National Commission for Astronomy.
\[ \hat{S} \text{ maximizes } P(Z \mid S)P(S). \]  
\[ \text{This is known as the maximum a posteriori (MAP) estimate of } S. \]

Equivalently, we can choose \( \hat{S} \) to minimize
\[ -\log P(Z \mid S) - \log P(S). \]  
\[ \text{The first term of } (3) \text{ will be familiar as the log likelihood of } S. \]

The second term can be thought of as a "roughness penalty," as images \( S \) which do not correspond to our prior conceptions will have been assigned small \( P(S) \) and hence a large penalty.

In statistical physics it is common to define probabilities by the "energy" \( U \) of a system, so that
\[ P(S) \propto \exp -\beta U(S), \]
\[ \text{where } \beta = 1/kT, \ T \text{ being temperature and } k \text{ Boltzmann's constant. If we adopt this notation, } S \text{ minimizes } \]
\[ -\log \text{ likelihood } + \beta U(S). \]

We can recognize this as a Lagrangian form, so its solution is equivalent to solving
\[ \text{max likelihood subject to energy } \leq \text{constraint} \]
and to
\[ \text{min energy subject to likelihood } \geq \text{constraint}. \]

Many other deconvolution principles fit into one of these forms, in particular maximum entropy (see, Narayan & Nitayananda 1986, for a review). What is important is that viewing \( P(S) \propto \exp [-\beta U(S)] \) as a prior distribution is a very illuminating way to select the energy function \( U \) and also an excellent way to consider how to choose parameters such as \( \beta \) as we will see later.

In maximum entropy methods \( U(S) = -\mathcal{H}, \) where \( \mathcal{H} = -\sum_p \log p_i \) is the Shannon entropy of the distribution \( (p_i) \) of photons to cells. That is, \( p_i \propto \kappa_i \) with \( \Sigma p_i = 1. \) This energy function does not depend on the spatial location of the pixels \( i \) in the image but only on the set of observed luminosities. In contrast, our energy function will be based on the spatial smoothness of the image \( S. \)

3. PRIOR MODEL

As we have seen in the previous section, the first ingredient in the Bayesian paradigm is the prior model. We will now describe the prior in our deconvolution problem.

The luminosity distribution of elliptical galaxies has been investigated by many astronomers, and several analytic functions have been proposed to model the distribution. Among others the most commonly used is the "r^{1/4} law" proposed by de Vaucouleurs,
\[ \log(I(r)/I_e) = -3.33 \left( r/r_e \right)^{1/4} - 1 \]
\[ \text{where } r \text{ is the distance from the center of the galaxy and } I_e \text{ and } r_e \text{ are parameters which differ from galaxy to galaxy. Furthermore, the luminosity distribution of pure disk in a galaxy can be modeled by the "exponential law" written as } \]
\[ I(r) = I(0) \exp (-b_0 r). \]

These results suggest that the luminosity of galaxies is most naturally considered on log scale, and we noted that astronomers tend to plot contour levels on galaxy images in a geometrical progression. We decided to model the smoothness of the luminosity distribution on log scale except for very small values of the luminosity.

For each pixel \( i \) let \( Y_i = \log(S_i + p) \) where \( p \) allows for a linear scale at very small photon counts (\( p = 100 \) in our examples). We will regard \( Y \) as a column vector of values of \( Y_i \), and similarly for \( S \) and \( Z \).

Our prior knowledge about the luminosity distribution makes it possible to model the distribution of \( Y \) by a conditional autoregression (see, Ripley 1981). We have
\[ P(Y) \propto \exp \left[ -\left( \frac{1}{2\sigma_x} \right) Y^T(I - \phi N)Y \right], \]
\[ \text{where } N_i = 1 \text{ if cells } i \text{ and } j \text{ are spatial neighbors (pixels at distance one), zero otherwise, and } \phi \text{ just less than 0.25. The term } Y^T(I - \phi N)Y \text{ represents in matrix notation the sum of squares of the values } Y_i \text{ minus } \phi \text{ times the sum of } Y_iY_j \text{ for neighboring pixels } i \text{ and } j. \]

The parameters can be interpreted by the following expressions describing the conditional distribution:
\[ E(Y_i \mid Y_{j \neq i}) = \phi \sum_{j \text{ nbr } i} Y_j, \]
\[ \text{var}(Y_i \mid Y_{j \neq i}) = \kappa_i, \]
where the suffix \( j \text{ nbr } i \) denotes the four neighbor pixels at distance one from pixel \( i \). The parameter \( \kappa_i \) measures the smoothness of the "true" image on log scale, and so is dimensionless. As it is a variance, we often quote its square root.

4. NOISE PROCESS

The observed image \( Z \) differs from the true brightness distribution \( S \) in having been blurred by atmospheric motion, and encountering statistical noise in the recording process. Both of these processes can be described quite precisely.

4.1 Blurring

A detailed review of the atmospheric processes which lead to the distortion has been presented by Woolf (1982; see also Young 1974). No exact expression describing the shape of the point spread function is known. However previous studies (Moffat 1969; Bouannou et al. 1983; Djorgovski 1983; Molina & Ripley 1989) have suggested a radially symmetric approximation for the point spread function \( h \) of the form
\[ h(r) \propto \left[ 1 + \left( r^2 / R^2 \right) \right]^{-\beta}. \]
\[ \text{This can be checked, and } \beta \text{ and } R \text{ chosen, by extracting what are clearly stars from a displayed image and fitting this function by weighted nonlinear least squares. The full width half at maximum (FWHM) for this function has the value } \]
\[ \text{FWHM} = 2R \sqrt{\beta} - 1. \]

For the images we analyze here we found \( \beta \gtrapprox 3 \) and \( R \gtrapprox 2.7 \) pixels which, with the scale of the telescope we use, corresponds to a FWHM \( \sim 1.4' \text{ to } 2.1' \). A point source is effectively spread out over around a hundred pixels. In Fig. 1 we plot the light profile of a star together with the fit to the proposed PSF as an example of the good fit of the analytical function.

4.2 Recording Noise

The principal sources of noises are the discrete nature of photons and electronic noise in the CCD detector. Physical considerations suggest that both sources will give noise which is independent from pixel to pixel with variance which may depend on the (average) number of photons arriving at the CCD cell. Let \( Z' \) denote the effect of blurring the true \( S \) by the point spread function \( h \). In vector notation \( Z' = HS \) where \( H \) is the matrix with entries \( H_{ij} = N \left[ \text{dist(pixel } i, \text{pixel } j \right] \). Then the photon count should
be Poisson with mean $Z_i$ and variance $Z_i$.  Adding electronic noise suggests

$$\text{var}(Z_i) = a + Z_i, \quad (15)$$

but since not all photons are recorded, we might expect

$$\text{var}(Z_i) = a + b Z_i, \quad (16)$$

since a (nonrandom) fraction $b < 1$ of photons are recorded, the number of arrivals has mean $Z_i/b$ with variance $Z_i/b$.  Thus $b$ times the number of actual arrivals has mean $Z_i$ and variance $b^2(Z_i/b) = bZ_i$.

Let $Z_i^n$ be the value at pixel $i$ minus the average value of its neighbors.  (We assume there are $r$ neighbors, four in our implementation.)  Then the mean of $Z_i^n$ is essentially zero, but the variance is $[1 + (1/r)]\sigma^2(\mu_i)$.  The idea of our estimation process is to divide the values of $\mu_i$ into small intervals, and to estimate $[1 + (1/r)]\sigma^2(\mu_i)$ from the variance of the $(Z_i^n)$ with $\mu_i$ in the interval.  There are two problems with this idea.  First we do not know $\mu_i$, only the observed value $Z_i$.  Fortunately, on the scale we are using, $Z_i$ has a mean of at least a few hundred and a variance about equal to its mean.  Thus the relative error in replacing $\mu_i$ by $Z_i$ is only a few percent at most.  The second problem is the occasional "gross errors" in the observations $Z_i$.  We found anomalous pixels whose value was tens of thousands different from the average of its neighbors, perhaps due to detector errors or cosmic-ray strikes.

Since we have so much data, we can afford to use a simple and inefficient method to estimate the variance while discarding "gross errors."  For observations $x_1, \ldots, x_n$ from a normal distribution with mean known to be zero but unknown variance $\sigma^2$ we would estimate $\sigma^2$ by

$$s^2 = \frac{1}{n} \sum x_i^2, \quad (17)$$

Since the observations are squared, this is extremely susceptible to outliers.  Consider instead an estimator proportional to $\Sigma|x_i|$ or median $|x_i|$.  Each of these would be expected to be proportional to $\sigma$, and much less susceptible to "gross errors," especially the second.  We use

$$\hat{\sigma}(\mu_i)^2 = \left[0.6745 \times \text{median} (|Z_i^n|)\right]^2 / (1 + 1/r), \quad (18)$$

where the median is over pixels with $Z_i$ in the range for $\mu_i$.  The constant $0.6745$ is that needed to obtain the correct value if there are no gross errors.

A little care is needed in interpreting $\hat{\sigma}(\mu)$.  This does not estimate $\text{var}(Z_i)$, which would be inflated by gross errors, but the variance without the errors.  This is correct, as our estimation procedure for $S$ uses robust statistics as described in Sec. 6 to remove gross errors.

An example of the estimation of $a$ and $b$ from a real galaxy is shown in Fig. 2(a).

5. DERIVATION OF THE MAP EQUATIONS

Having defined the prior model and the noise process for our deconvolution problem we may estimate the true underlying image.  The maximum a posteriori (MAP) estimator is that "true image" $S$ which minimizes

$$-2 \log P(S | Z) = \text{const} - 2 \log P(Z | S) - 2 \log P(S). \quad (19)$$

As we saw in Sec. 3 we work not with $S$ but with the image $Y$, where $Y_i = \log(S_i + p)$ for each pixel.  Thus we choose to find the MAP estimate of $Y$, not of $S$.  (It would be desirable if $S_i = \exp(Y_i) - p$ gave the MAP estimate of $S$, but this is not the case.  Since we must choose, we give $Y$ higher priority.)  Thus we choose $Y$ to minimize

$$-2 \log P(Z | S) - 2 \log P(Y), \quad (20)$$

since $P(Z | S) = P(Z | Y)$.  Under our noise model at each pixel $i$ the observed value is Gaussian, with mean $(HS)_i$, and variance $\sigma^2(HS)_i$.  Thus

$$-2 \log P(Z | S) = \text{const} + \sum_{p=1}^{n} \left[ (Z_i - (HS)_i)^2 / \sigma^2(HS)_i \right]$$

$$- \log \sigma^2(HS)_i, \quad (21)$$

and for our prior

$$-2 \log P(Y) = 2 U(Y) = (1/\kappa) Y^T(I - \phi N) Y. \quad (22)$$

Differentiating (20) with respect to each $Y_i$ in turn we find that either $S_i = 0$ or

$$Y_i = \left[ \phi N Y + \kappa, \text{diag} (e^T) H^T W^{-1} (Z - HS) \right], \quad (23)$$

Fig. 2.  (a) Robust estimator of $\sigma^2(Z)$ vs $Z$ and the fit to a straight line for NGC 7320.  (b) $\kappa$, estimation for NGC 450/UGC 807.
where \( W = \text{diag} \{ \sigma^2(HS_\lambda) \} \) is a diagonal matrix of variances. In this process we ignore the dependence of \( \sigma^2(HS_\lambda) \) on \( Y \); in Molina & Ripley (1989) we show that this approximation has negligible effect. Combining these pieces, we find that \( \hat{Y} \), the MAP estimator of \( Y \), satisfies

\[
\hat{Y} = \max \{ \log(p) I, \phi N + \kappa, \text{diag}(e^\delta)H^TW^{-1}(Z - H\hat{S}) \},
\]

\[
\hat{S} = \exp(\hat{Y}) - p.
\]

These equations are solved by iteration. Suppose that \( Y^{(i)} \) is the current image. Forming \( S^{(i)} = \exp(\hat{Y}^{(i)}) - p \) and substituting \( Y^{(i)} \) and \( S^{(i)} \) on the right-hand side of (24) gives \( Y' \), say. To stabilize the convergence procedure we only move part of the way from \( Y^{(i)} \) to \( Y' \). For each pixel \( i \) we have

\[
Y'^{(i+1)} = \max \{ Y'^{(i)} - \log(2),
\min \{ Y'^{(i)} + \log(2), cY' + (1-c)Y^{(i)} \} \},
\]

where \( c \) is a small positive constant. This shows that the value at each pixel moves only a small proportion of the way to \( Y' \), and the new value of \( S_i \) differs from the old by at most a factor of 2.

6. ROBUST FITTING

Our early experiments with iterations using (24,25) had considerable difficulties near anomalous pixels. The methods of robust statistics (Huber 1981; Hampel et al. 1986) were developed to deal with precisely such problems, called “gross errors” in the robust statistics literature. The idea is to downweight observations which are sufficiently far away from their mean. Such values are given too much weight by the first term of (21). The squared term in \( D_i = [2(Z_i - H\lambda)]^2/\sigma^2(HS_\lambda) \) represents the number of standard deviations that \( Z_i \) is away from its mean. In robust statistics we replace \( \Sigma D_i^2 \) by \( \Sigma \rho(|D_i|) \) for a function \( \rho \) which penalizes extreme values less severely.

We used Huber’s “proposal 2” function for \( \rho \), defined by

\[
\rho(x) = \begin{cases} 
  x^2 & \text{for } |x| \leq h, \\
  2h |x| - h^2 & \text{for } |x| > h.
\end{cases}
\]

This is quadratic in the center, but penalizes large deviations linearly rather than quadratically. Equivalently, observations \( D_i \) are downweighted if \( |D_i| \) exceeds \( h \). In practice we choose \( h \) about 2, so that only observations more than two standard deviations from the mean are downweighted.

Let \( \psi \) denote the derivative of \( \rho \). For our choice \( \psi(x) = x \) for \( |x| < h \), and \( \psi(x) = h \text{sgn}(x) \) for \( |x| > h \). The robust MAP estimator of \( Y \) satisfies (Molina & Ripley 1989)

\[
Y = \max \{ \log(p) I, \phi N + \kappa, \text{diag}(e^\delta) H^TW^{-1/2} \psi( W^{1/2}(Z - HS) ) \},
\]

The effect of the function \( \psi \) is to reduce the size of terms with anomalously large or small \( Z_i \) to a manageable level, about the size of a large chance deviation.

7. PARAMETER ESTIMATION

The parameter \( \kappa \), determines how smooth the image is expected to be (on log scale) and hence how smooth the restorations will be. It would be expected to be essentially constant under fixed viewing conditions (telescope and detector), although we might expect elliptical galaxies, for example, to be smoother than spiral galaxies.

The method we use to estimate \( \kappa \) is based on the idea that high-frequency information in \( S \) is lost by the blurring and masked by the addition of noise, but the low-frequency information can be used directly to estimate \( \kappa \). That is, we consider \( \log(Z + p) \) at low frequencies as if it were \( Y = \log(S + p) \). Since \( \kappa_n \) is a measure of the power of \( Y \), we estimate it by measuring the low-frequency power of \( \log(Z + p) \).

More formally we know that

\[
Z = H(e^Y - pI) + \text{noise} = He^Y - pI + \text{noise}.
\]

The matrix \( H \) operates locally, where values of \( Y \) are relatively constant, so to a crude approximation

\[
LZ = \log(Z + p) \approx HY,
\]

with the approximation being best at low frequencies.

If \( Y \) were observable, the maximum likelihood estimator of \( \kappa_n \) would be \( Y^T(I - \phi N)/n \), where \( n \) is the number of pixels. Using \( \gamma(\omega) \) and so on to indicate (two-dimensional) Fourier transforms, this is \( \Sigma [1 - \phi n(\omega_i)] |\gamma(\omega_i)|^2 \) where the sum is over Fourier frequencies. Furthermore, we find that, for any set \( D \) of Fourier frequencies,

\[
\Sigma_D [1 - \phi n(\omega_i)] |h(\omega_i)|^2
\]

\[
= \Sigma_D [1 - \phi n(\omega_i)] |h(\omega_i)|^2 |\gamma(\omega_i)|^2
\]

\[(30)\]

has mean

\[
\Sigma_D h(\omega_i)^2 \kappa_i.
\]

(31)

Substituting (29) into (30) and using a set \( D \) of low frequencies we find that

\[
\Sigma_D [1 - \phi n(\omega_i)] |\tau(\omega_i)|^2 / \Sigma_D |h(\omega_i)|^2
\]

\[(32)\]

is an approximately unbiased estimator of \( \kappa_i \). The close relationship to the maximum likelihood estimator suggests that it should be an efficient estimator; this is borne out by our experience.

Figure 2(b) shows the results of our estimator \( \hat{\kappa}_i \) for the NGC 450/UGC 807 image (see Sec. 10). We found that we were always able to use a fair proportion of the frequency range, and that the exact cutoff value is not at all critical. Past experience suggests method for choosing where the \( \kappa \) vs frequency function starts to be almost flat, in particular, for our test images we obtain \( \kappa_i^{1/2} \approx 0.07 - 0.15 \).
8. CHOICE OF INITIAL SOLUTION

The non-negativity of $S$ is enforced by Eqs. (24), and it is a very good idea to choose a starting value $S^{(0)}$ which satisfies this condition. One possible choice is

$$S^{(0)} = \max(Y_i, Z_i).$$

(33)

This will usually be a long way from the solution, especially at the galaxy peaks. This forces us to choose the relaxation constant $c$ in (25) to be small, and hence convergence will be slow. (Often $c$ can be increased as the process nears convergence.) If the computational time available is this is a perfectly feasible procedure—we have found 40–200 iterations to be necessary.

The computational load can be reduced by choosing a starting point closer to the expected solution. This can be done by a series of approximations. Our prior distribution for $Y$ is such that conditional on its neighbors $Y_i$ has mean $(\phi N Y)$, and variance $\kappa_i$. Then, by a local linear expansion

$$E(S_i | S_j \neq i) \approx (\phi N S),$$

$$\text{var}(S_i | S_j \neq i) \approx \kappa_i [E(S_i) + p].$$

Therefore, locally, we can pretend we have a CAR model for $S$ with variance $\kappa_i [E(S_i) + p]$, and a noise model with variance $\sigma^2[HS] = a + b(\text{HS})$. For this system the MAP equations (Molina & Ripley 1989) are

$$(H^T H + \lambda I - \lambda \phi N) S = H^T Z$$

(34)

where

$$\lambda = \frac{a + b(\text{HS})}{\kappa_i [E(S_i) + p]}$$

(35)

and in regions where $S_i$ is large, $\lambda \approx b / \kappa_i$. We conclude that (34) with this value of $\lambda$ is a good approximation at least in the regions of the image of high luminosity. The great advantage of (34) is that it can be solved directly by Fourier methods. If $h(\omega)$ and so on once again represent Fourier transforms, (34) is equivalent to

$$[|h(\omega)|^2 + \lambda I - \lambda \phi n(\omega)]s(\omega) = h(\omega)z(\omega),$$

and since $N$ is symmetric the expression [...] is real and $s(\omega)$ is obtained from $z(\omega)$ by a linear filter. This is a classical method of regularization in deconvolution. We use it to find a starting point $S^{(0)}$. Note that positivity is not guaranteed, but can be forced as at (33). With this starting point many fewer iterations are needed, perhaps 10–20.

9. COMPUTATIONAL EXPERIENCE

Fast Fourier transforms are used both to compute the starting point via (8) and to compute the right-hand side of (24), in that the convolutions implied by $H$ and $H^T$ are done in frequency domain. Efficient FFTs are essential to the efficient implementation of our method. After some experiment with the base 2 and 4 routines of Singleton (1969), we found also to make use of the fact that $Z, S,$ and $h(\omega)$ are all real and use real FFTs wherever possible. Thus, this method is seen to be less than it would be; it is explained in detail in Press et al. (1986). The use of Fourier transforms necessitates the extension of the images to sizes which are powers of two such as $512^2$.

Our implementations have been on a VAX 6320 and on Sun 3 and 4 workstations. One iteration on a $512^2$ images takes 72 s on a VAX 6320 and 27 s on a Sun 4/370. The deconvolution of such an image takes about ten minutes on the Sun.

10. TEST EXAMPLES

This aim of this section is to describe how our Bayesian model for image analysis works on synthetic and real astronomical objects. These objects have been chosen as to cover the widest possible range of light distributions that normally appear in extended astronomical objects.

10.1 Applications to CCD Images of Galaxies

Several CCD frames of galaxies were analyzed to test the method with real data. We will discuss the results on two very different types of galaxies, the analyses of an elliptical (NGC 7675) and of a Sc spiral galaxy, in order to test the method against two very different luminosity distributions and the presence of small scale features such as H II regions. Table 1 summarizes the set of observational data we used.

The images were obtained at the prime focus of the 3.5 m telescope of the Centro Hispano-Alemán in Calar Alto (Almeria, Spain) and at the Cassegrain focus of the JKT 1 m telescope in the Observatorio del Roque de los Muchachos (La Palma, Spain).

10.1.1 NGC 7675

This object is an early type galaxy member of group 96 of the Hickson’s (1982) compilation of compact groups. Its morphological classification corresponds to E2 as denoted by Hickson et al. (1986). In Fig. 3 we show the original

<table>
<thead>
<tr>
<th>Galaxy</th>
<th>Band</th>
<th>Size</th>
<th>Telescope</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGC 7675</td>
<td>F</td>
<td>256×512</td>
<td>3.5</td>
</tr>
<tr>
<td>NGC 450/UGC 807</td>
<td>F</td>
<td>512×512</td>
<td>3.5</td>
</tr>
<tr>
<td>NGC 7320</td>
<td>F</td>
<td>512×512</td>
<td>3.5</td>
</tr>
</tbody>
</table>

© American Astronomical Society • Provided by the NASA Astrophysics Data System
frame together with three restorations for different values of the relaxation parameter $c$ and $\kappa$. As can be seen from the figure a good compromise for the restoration is a deconvolution with values $c \approx 0.1$ and $\kappa^{1/2} \approx 0.1$ which results in a sharp restoration with a reasonable degree of smoothness for the large scale features. Note that the restoration is not sensitive to the relaxation factor $c$ at all, and this is a good indication that the iterative process has converged correctly. In order to estimate some observational profiles, sets of isophotes were selected for the original and deconvolved images and were fitted by ellipses of arbitrary minor axis, ellipticity, and orientation. In Fig. 4 we present the surface brightness as a function of the semi-minor axis for the two images. As can be seen clearly from the figure, the profile of the deconvolved image appears to be very well represented by an $r^{1/4}$ law, a fact which cannot be deduced from the original image which has a core structure due to blurring.

Blurring effects appear clearly when analyzing isophote twisting or changes in ellipticity. In Figs. 5(a) and 5(b), we present the variations of the orientation and the ellipticity, respectively, for the original and deconvolved images. As a result of the deconvolution process the intermediate and small scale spatial behavior of the two quantities may be recovered, suggesting the triaxial nature of NGC 7675.

10.1.2 NGC 450/UGC 807

This is a case of two galaxies having a small angular separation but with very discrepant redshifts. Several deconvolutions were performed with a spread of values for the $\kappa$ parameter. A good compromise between the degree of smoothness and resolution is found to be $\kappa^{1/2} \approx 0.07$ and that value is used here as representative for the restoration of this image. The original and deconvolved frames are presented in Fig. 6 [Plate 48]. As can be seen from the figure many small scale features are strongly enhanced mainly at the disk of NGC 450 and the arms of its companion are now clearly delimited.

In order to identify small scale features in NGC 450 an oversmooth ($\kappa^{1/2} = 0.007$) restoration was subtracted from the selected deconvolution and the result was compared with a redshifted H\textalpha frame. In Fig. 7 we present a grey-scale plot of the subtraction with H\textalpha contours superimposed onto it. The agreement between the positions of local maxima between both images indicate they can be identified as H\textalpha.
regions whose light is dominated in the V band by the Hβ and [O III] (5007 Å) emission lines. An internal structure in the bulge of the galaxy appears in the deconvolved image. This feature is not detected in the original image but is present in the Hz frame.

These comparisons with the Hz frame do indicate that the additional detail produced by deconvolution is genuine.

10.1.3 Additional examples

Galaxies a and b of Hickson’s group 92 (Stephan’s Quintet) were analyzed. In Figs. 8 and 9 we present contour plots for these two galaxies together with the deconvolutions performed with a value $\kappa_{1/2} = 0.15$. Galaxy a (NGC 7320) is a Sd spiral with several star forming regions which are clearly delimited when applying the method and in excellent agreement with the results already found by Arp (1973) who presented a detailed list of H II regions for this galaxy on the basis of Hz surface photometry. Some ringing effects appear on a bright star superposed on the galaxy but this is due to saturation in the original CCD frame.

10.2 Synthetic Images

Three synthetic 256 x 256 images are presented to test the method. These represent an exponential disk, an exponential disk plus five Gaussian knots, and an $r^{1/4}$ galaxy. In Fig. 10 we present the results of the method when applied to a disk of central flux $I(0) = 10^5$ counts and $b_0 = 0.031$ pixels$^{-1}$ convolved with a PSF with $R = 6$ pixels (which with the scale of our telescopes would correspond to about 2”). In this case we have added signal dependent Gaussian noise with corresponding values $a = 200$ and $b = 0.2$ [see Eq. (16)]. The estimated values we obtained were $a = 112$ and $b = 0.25$, but removing just one outlier of the points shown in Fig. 11(a) we obtain $a = 140$ and $b = 0.2$. In Fig. 11(b) we present a plot of the estimation of the $\kappa$ parameter for this image. We used $\kappa_{1/2} = 0.4$ obtaining the deconvolution in 100 sweeps with relaxation parameter $c = 0.1$.

In Fig. 12 we present a similar analysis of an image which consists in the disk of the previous example plus five Gaussians with a central flux of $2 \times 10^6$ counts and $\sigma = 3$ pixels. In this case the image to be deconvolved was created with $R = 3$ pixels, $a = 50$ and $b = 0.1$, and the deconvolution was performed with $\kappa_{1/2} = 0.1$ and $c = 0.1$.

As can be seen from both examples, a proper deconvolution can be obtained under several observation conditions and under different levels of noise. The method seem to be well suited for exponential profiles.

The last synthetic example consists of an $r^{1/4}$ elliptical galaxy with a constant axial ratio of 0.6. In this case the convolution was performed with $R = 4$ pixels and $a = 50$, $b = 0.2$, and the restoration was obtained using $\kappa_{1/2} = 0.06$, $c = 0.1$. In Fig. 13 we present plots of the same contour levels for the original, convolved plus noise, and restored images.
Fig. 10. Forty degrees perspective for an exponential disk, the convolved image plus noise and the restoration.

Fig. 11. (a) Same as Fig. 2(a) for the first synthetic image. (b) Same as Fig. 2(b) for the first synthetic image.
Fig. 12. Contour levels of a synthetic exponential disk plus five Gaussian knots, the convolved image plus noise and the restoration. Levels are $20 \times 2^k$ ($k = 1, \ldots, 13$). Also presented is a cross section for the three images.

Fig. 13. Contour plots of a synthetic $r^{1/4}$ elliptical galaxy (a), the convolved and noisy image (b), and the restoration (c).
14 a plot of the axial ratio as a function of the radius is displayed. The restored image indicates much more clearly the $r^{1/4}$ nature of the galaxy.

In all the examples we have found that the method allows for variations for each parameter by at least a factor of two or three, obtaining essentially the same restorations. With this kind of deconvolution technique a set of restorations can be easily built and compared as the basis of other analysis (profiles, ellipticity, twisting, presence of small features as H II regions or globular clusters, multiband images, etc.).

11. CONCLUSIONS

A novel Bayesian method to correct for the effect of the seeing in the presence of noise is presented. The method was tested with long exposure CCD frames of extragalactic objects. We made a short preliminary analysis of galaxies with different morphological types showing the utility of such techniques.

For the case of NGC 7675 we found that its surface brightness profiles may be well represented by a de Vaucouleurs' law, a fact which cannot be deduced directly from the original data. The effect of the blurring on the variation of the position angle and of the ellipticity with the minor axis is found to be even more important than in the light profile, indicating that this method is well suited to analyze such effects in the context of early type galaxies.

A restoration was also performed for an image of the pair NGC 450/UGC 807, two spiral galaxies. In this case the H$\alpha$ light distribution was used to test if small scale features such as H II regions can be identified in broadband photometry when removing the effects of the seeing and of the noise of the observed images. The results indicate that this is the case and with some simple image processing (removing low frequency patterns) the very small structures in the bulges of spirals can be recognized within scales of typically 1–3 pixels (0.3'–0.9') with the scale of the telescope we used.

Finally, one of the major advantages of this Bayesian method is that it is able to work with images with as large a number of pixels as our present detectors. In fact, with the CPU times quoted previously, the method takes less than an hour to deconvolve a $512 \times 512$ image, being much faster than many other alternative methods.

We thank Dr. M. Moles and Dr. V. Costa for providing us with the H$\alpha$ frame of NGC 450. This work was partially supported by the British Council, the UK Science and Engineering Research Council, and by a grant from DGICYT, reference PB87-0251. One of us (A.O.) acknowledges a grant from the Spanish Consejo Superior de Investigaciones Científicas. Dr. A. I. Sutherland provided computational assistance.

REFERENCES

Franken, S., & Thomsen, B. 1980, in ESO Workshop on Two-Dimensional Photometry, edited by P. Crane and K. Kjær (ESO, Munich, FRG), p. 319
Singleton, R. C. 1969, An algorithm for computing the mixed radii fast Fourier, IEEE Transactions on Audio and Electroacoustics, AU-17, pp. 93–103
NGC 450/UGC 807

\[ \kappa_s^{1/2} = 0.07 \]

Original

Deconvolution

Fig. 6. Original and deconvolved images of NGC 450/UGC 807.

Molina et al. (see page 671)