

Scale sensitive deconvolution

S. Bhatnagar and T.J. Cornwell

National Radio Astronomy Observatory (NRAO), Socorro, NM - 87801
Emails: sbhatnag@aoc.nrao.edu, tcornwel@aoc.nrao.edu

Abstract. Aperture synthesis radio telescopes measure the Fourier transform of the sky brightness distribution. However the Point Spread Function (PSF) of such telescopes has significant and widespread side-lobes, which needs to be deconvolved from the images. Existing deconvolution algorithms can be thought of as decomposing the image into a set of delta functions (scale less basis). This uses more degrees of freedom than necessary and is not optimal for extended emission. In this paper we present an iterative scale sensitive deconvolution algorithm for radio interferometric imaging, which attempts to minimize the degrees of freedom used to represent the signal (spatially correlated pixels).

1. Introduction

A radio interferometric telescope incompletely measures the visibility function, at discrete points. The Fourier transform of the visibility function, called the Dirty Image (I^D), is the convolution of the true image (I^o) with the telescope Point Spread Function (PSF)

$$I^D = \mathcal{FT}(V^o S) = I^o \star B \quad (1)$$

where (B) is the PSF, S the visibility sampling function and V^o is the observed visibility.

The goal of deconvolution algorithms is to estimate a sky image model I^M , such that the model visibility $V^M = \mathcal{FT}^{-1}(I^M \star B)$ fits the observed visibility to the extent allowed by the noise. A generalized model image I^M can be expressed as a linear sum of Pixel models

$$I^M = \sum_{k=0}^{N_M} P_m(\vec{p}_k) \quad (2)$$

where $P(\vec{p}_k)$ is the pixel model, \vec{p} are the parameters for the amplitude, location and the shape of P . The problem of optimal deconvolution then reduces to solving for I^M with a minimal set $\{\vec{p}\}$ allowed by the data.

The current popular image deconvolution algorithms (Karovska, 2002), like CLEAN (and its variants (Clark, 1980; Cornwell et al., 1990) and MEM (and its variants (Cornwell&Evans, 1985)) model I^o in a scale-less basis (delta functions). Such algorithms also require regularizes to avoid over-fitting (which results into

spurious compact sources in the image). Usually, this regularization is done via a user defined maximum number of components or/and global estimate of the noise in the image. Extended emission, which is at a very different scale than a compact component, is broken up into delta functions and later smoothed to suppress the high frequency errors made in such a representation. However since delta functions are at a scale smaller than even the resolution element, this results into the well known breaking-up of extended emission problem. In this paper we describe an algorithm which decomposes the sky image into parameterized Adaptive Scale Pixel (Asp) model. The parameters of the Aspen are determined using non-linear minimization. The algorithm is sensitive to the local spatial scale as well as the local signal-to-noise ratio.

2. Algorithm

The functional form for the Asp used in this paper is a symmetric two dimensional Gaussian. The algorithm searches for the locally best fit Asp to the Dirty Image, by estimating the location (x_k, y_k) , amplitude (A_k) and the size (σ_k) of the Asp.

The dirty image is smoothed to a few scales ranging from the smallest to the largest expected scale. A global set of Aspen, $\{P_o\}$ is maintained and a new Asp added to this list at each iteration. The model image is computed using this set and Eq. 2. The image decomposition into Aspen basis then proceeds as follows:

1. At each iteration, compute the model and residual visibilities as $V_i^R = V_i^M - V^o$, where V^o is the observed visibilities. Compute the residual image $I^R = \mathcal{FT}(V^R)$ and smoothed versions at a few scales.
2. Locate the peak among all the smoothed versions of I^R . Use the location, amplitude and the scale at which the peak was found as the initial guess for the new Asp P_k . Set $\{P_o\} = \{P_o\} + P_k$.
3. Make a sub-set of Aspen $\{P_i\}$ which will maximally affect the convergence.
4. Simultaneously solve for the parameters of this sub-set such that $\chi^2 = \sum |V^o - V_i^M|^2$ is minimized.
5. Goto Step 1 till I^R is noise like.

Ideally, all the Aspen determined in the earlier iterations should be kept in the problem at each iteration. However since the computational load scales strongly with the number of Aspen in the problem, the speed of convergence is significantly improved by adaptively dropping those Aspen which may not affect the χ^2 at the current iteration. Since the scale of the local Aspen is also adaptively determined based on the local signal-to-noise ratio (SNR) and the active set of Aspen determined in previous iterations are kept in the problem, the number of effective parameters used to represent the final image is minimized.

The set of active Aspen is determined by applying a threshold on the length of the vector of the first derivatives of the χ^2 with respect to all the parameters (p_i) of each Aspen P_k $D_k = \sum_i [(\partial\chi^2/\partial p_i)^2]^{1/2}$. D_k is computed for each k at the start of each iteration and all Aspen with $D_k < D_o$ are dropped from the problem at that iteration. Since this is done at the beginning of each cycle, mistakes in this estimate for the active set of Aspen are corrected in later iterations. This

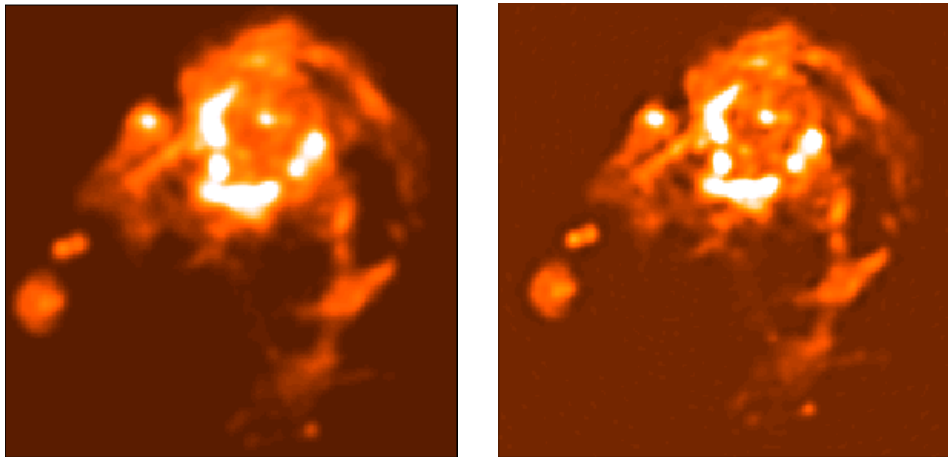


Figure 1. Figure showing the results of the Asp deconvolution on a test image. The left panel shows the original model image used. A 600 Asp model of the image after deconvolving the PSF is shown in the right panel. The scale of the right image is adjusted to show the low level noise as well.

however implicitly assumes that the χ^2 surface has the same curvature along all axes. Ideally, the active set should be determined by thresholding the covariance matrix, which is computationally expensive. Since the value of the off-diagonal elements and the structure of the covariance matrix is strongly dependent on the side-lobe levels of the PSF, the active set of Aspen can be estimated by using an Aspen decomposition of the PSF and its significant side-lobes. Such a PSF decomposition (Bhatnagar&Cornwell, 2003) appears to model the distant but significant side-lobes of a typical PSF well. Use of such an approximation for the PSF to compute approximate covariance matrix is being currently investigated.

3. Results

Several tests were done using simulated model images (not shown here) which were convolved with PSF for typical VLA observations and then deconvolved using the above algorithm. The algorithm was found to be sensitive to local scales and even overlapping components with varied scales were separately detected. Consequently, the model image was represented with many fewer degrees of freedom compared to other scale-less or multi-scale algorithms (Holdaway&Cornwell, 2001).

A VLA observation of M31 was used as a more realistic test case. The best available deconvolved image was used as the “ideal” model image. The model image was then convolved with a PSF corresponding to a typical VLA C-array observation. The resulting dirty image was then deconvolved using the above algorithm. The original model image, and the model image from this algorithm are shown in Fig 1. The Asp model in the right panel is composed of 600 Asp. For similar dynamic range, the multi-scale (MS) Clean used ~ 8000 components. The scale-less Clean algorithm required about a factor 10 more components than

even MS-Clean, and did not do as well as Asp or MS-Clean in terms of residual noise and fidelity.

4. Conclusions

Image deconvolution can be treated as a search for a model image, which when convolved with the PSF, minimizes an objective function (e.g. the χ^2). The model image is represented as a parameterized function, which is estimated by the deconvolution algorithms. Since the residual image provides the update direction in an iterative scheme, a parameterization which fundamentally separates noise from the signal (the sky emission) will always produce better results in terms of dynamic range and fidelity. Correlation length (or equivalently the scale of emission) is one such parameter which strongly separates the noise (which is fundamentally scale-less) from the signal. The algorithm presented here uses scale as one of the parameters and given the PSF, attempts to separate noise and signal using the Adaptive Scale Pixel (Asp) model. It is sensitive to the local scale of emission and SNR and is shown to perform better even for complex images with a range of scales. The reconstruction is optimal at all scales using minimum degrees of freedom compared to other algorithms. Heuristics used to eliminate insignificant Aspen, which adaptively changes the dimensionality of the problem at each iteration, are shown to be effective. Other, possibly more effective methods for this using an estimate of the covariance matrix are being explored. Also, a pixel model with a tighter support constraint, with also convenient mathematical properties is also likely to improve speed of convergence. Work is in progress to incorporate these improvements in the algorithm.

Acknowledgments. We thank R.V.Urvashi, Divya Oberoi and Jeff Kern for useful discussions and comments.

References

- Bhatnagar, S. & Cornwell, T.J., 2003, Proc. of the Annual SPIE Meeting, In press
Clark, B.G. 1980, A&A, 89, 377
Cornwell, T.J., Evans, K.F. 1985, A&A, 143, 77–83
Cornwell, T.J., Braun, R., & Briggs, D.S. 1990, ASP
Holdaway, M.H., Cornwell, T.J. 2001, in preparation
Karovska, M., 2002 BAAS, 201, 6302