DIRECT MAXIMUM-ENTROPY IMAGE RECONSTRUCTION FROM THE BISPECTRUM

D.F. BUSCHER*

NRL/USNO Optical Interferometer Project, Washington DC

1. Introduction

The strategies for image reconstruction from optical synthetic-aperture data can be divided into two camps according to their historical legacies: those coming from the field of more conventional optical image processing, where the data are relatively complete in terms of Fourier coverage, have concentrated on trying to recover images using "brute force" methods. They make use of the massive size of the typical datasets to try and derive an image without having to use any a-priori constraints. The size of the datasets to some extent restricts the algorithms used towards being as simple and hence fast as possible. Radio-astronomical imaging strategies, on the other hand, have always been designed to cope with sparse and missing data, and many sophisticated algorithms have been designed to make maximum use of a-priori constraints with iterative global fitting techniques.

However, existing radio-astronomical imaging packages do not make use of the extra redundancy in optical data, nor do they use the correct model for the noise experienced in optical aperture synthesis measurements. It should be noted that this is not because of the quantum-limited nature of optical interferometric measurements: it is due to the fact that optical datasets typically consist of incoherent averages of many thousands of exposures, while in radio VLBI little or no incoherent averaging is performed. The result of this difference in measurement strategy is that in optical interferometry, the primary observables are the averaged visibility squared (or power spectrum) and the averaged triple product (or bispectrum) whereas the primary observable in radio VLBI is the complex visibility, albeit corrupted by antenna-gain errors. Thus, while radio-astronomical software has been used successfully in optical interferometry (see e.g. Haniff *et al.* 1989), these algorithms do not always make the best use of the data.

I present here an imaging package specifically intended for optical aperture synthesis work, but which adopts a philosophy derived from radio aperture synthesis. This approach adds flexibility and robustness to the image reconstruction process, and is not as computationally demanding as might at first be thought.

2. Algorithm

The algorithm used results straightforwardly from the problem statement "What images are consistent with the available bispectrum and power spectrum data (and any a-priori image constraints)?" Since the bispectrum and power spectrum data are the result of averaging many interferograms, the noise on the data can reasonably

---

* Also at Universities Space Research Association, Washington DC

be assumed to be Gaussian, and thus the question of consistency between a given image and the data can be derived from the familiar χ^2 statistic. Roughly speaking, any image whose data residuals satisfy $\chi^2 \sim N$, where N is the number of data measurements, is consistent with (i.e. "fits") the data. Our a-priori knowledge allows us to reject all images which are not positive everywhere, or contain flux outside some window but, as with any noisy measurement, we will usually be left with an infinite number of possible images which are consistent with the data. We have chosen here to use the Maximum-Entropy criterion to "regularize" the problem, i.e. we choose that image which fits the data and also has maximum configurational entropy.

To find this chosen image, one could imagine generating random images, determining for each image the data measurements such an image would have produced, computing the χ^2 and entropy values, and selecting the image which satisfies all the constraints. While such a scheme is inefficient, it has the advantage of not requiring any mathematics beyond the writing down of the measurement equation that derives mock data from a test image: no computation of the inverse transformation is needed. In cases such as the bispectrum, where no general inverse exists, this advantage is significant. Furthermore, because there is no explicit inverse equation, this scheme will work on any subset of the data - there is no requirement for completeness in any sense. Of course, incomplete data will imperfectly constrain the image, but will not invalidate our question "which images fit the data?" Finally, we can be assured that the image finally selected explicitly satisfies all our constraints: no error-propagation calculations are needed, since the image is a result of the global fit of the image to the data — no intermediate steps are involved.

The importance of this last point cannot be overemphasized: if we have found an image which satisfies all our constraints, then by definition no other image can be said to be "better", unless new data or a-priori criteria are introduced.

A more efficient approach which retains the above advantages would be to do a directed search, i.e. to make use of information derived from previous images to determine which new images to try. Such an approach is represented by the "building-block" method (Hoffman & Weigelt 1991). The method used here makes use of gradient-descent methods to further increase the efficiency of the search. This involves the extra step of calculating the derivatives of the measurement equation, but these calculations are in fact surprisingly "cheap" computationally (unfortunately space does not allow derivation of this result here).

The program presented here uses the Maximum-Entropy library MEMSYS2 (Maximum Entropy Data Consultants, Cambridge, UK) to implement a gradient-descent algorithm encorporating the Maximum Entropy criterion. The program is relatively simple and fast: a 32×32 pixel image can be reconstructed from 5000 data points in under 3 minutes on a 25MHz 386 IBM PC.

3. Results

Space suffices to show only a few preliminary reconstructions. Figure 1 shows images reconstructed from simulated data with signal-to-noise ratios (SNRs) varying from 5.0 down to 0.16. It can be seen that the image reconstruction process suc-

ceeds at lower SNRs than the value of 2 commonly quoted as the lower limit for other non-linear image reconstruction techniques such as self-calibration. This may not seem surprising since a large number of bispectrum measurements were used in the reconstruction, but the conventional wisdom states that non-linear algorithms behave differently from linear algorithms in that a large amount of data cannot compensate for a lower SNR beyond some lower limit in SNR. The only way to make use of low SNR measurements is to combine the measurements linearly, e.g. by averaging; this was not done here as there is no general linear relationship between the bispectrum measurements. The phenomenon of a SNR threshold for reconstruction still occurs: when the simulations were repeated with 4 times the amount of bispectrum data, the reconstructions began to fail at similar SNRs. What appears to be happening is that the SNR threshold for bispectrum reconstruction is lower than for self-calibration, perhaps because there is some implicit linear combination of bispectrum data taking place (Cornwell 1993, private communication).
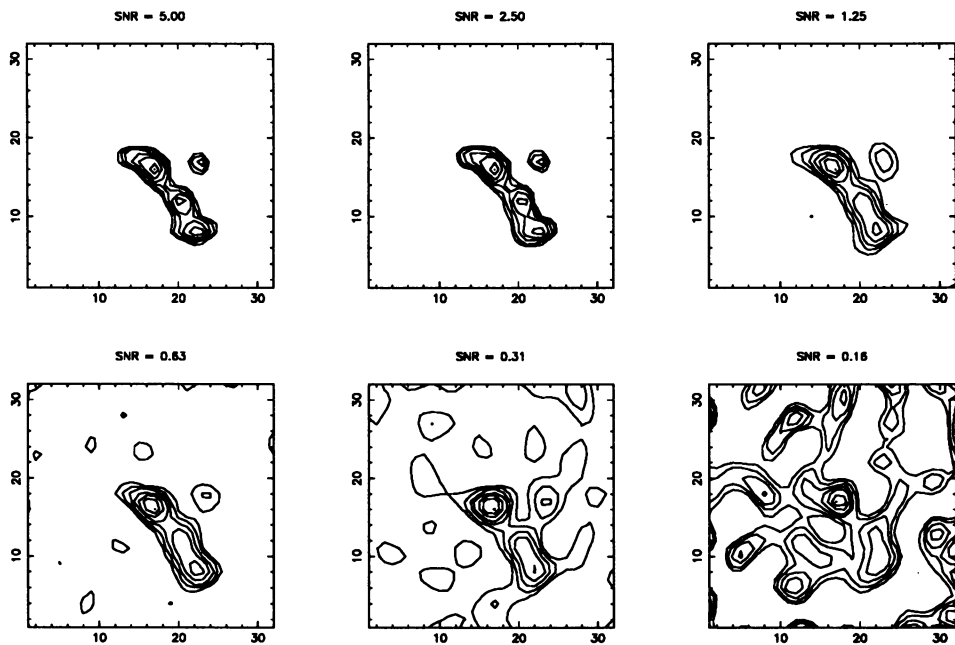
Fig. 1: Image reconstructions from simulated data. The SNR of the data is indicated at the top of each image. The SNR is uniform across each dataset, comprising 150 power spectrum points and 5,000 bispectrum points. The contour levels in all the images are at 2%, 4%, 8%, 16%, 32%, 64% and 99% of the peak intensity.

References

Haniff, C.A., Buscher, D.F., Christou, J.C., & Ridgway, S.T., 1989. "Synthetic aperture imaging at infrared wavelengths" *M. Not. R. Astr. Soc.*, **241**, 51P–56P.

Hofmann, K.-H. and Weigelt, G., 1991. 'The building block method: image reconstruction from the bispectrum using an iterative algorithm', in: *High Resolution Imaging by Interferometry II*, ed. F. Merkle, ESO: Garching bei München.