Source Detection with Interferometric Datasets

Cathryn M. Trott[1,2], Randall B. Wayth[1], Jean-Pierre R. Macquart[1], and Steven J. Tingay[1]

[1]International Centre for Radio Astronomy Research, Curtin University, WA 6845, Australia
email: cathryn.trott@curtin.edu.au
[2]ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO)

Abstract. The detection of sources in interferometric radio data typically relies on extracting information from images, formed by Fourier transform of the underlying visibility dataset, and CLEANed of contaminating sidelobes through iterative deconvolution. Variable and transient radio sources span a large range of variability timescales, and their study has the potential to enhance our knowledge of the dynamic universe. Their detection and classification involve large data rates and non-stationary PSFs, commensal observing programs and ambitious science goals, and will demand a paradigm shift in the deployment of next-generation instruments. Optimal source detection and classification in real time requires efficient and automated algorithms. On short time-scales variability can be probed with an optimal matched filter detector applied directly to the visibility dataset. This paper shows the design of such a detector, and some preliminary detection performance results.

Keywords. methods: statistical, techniques: interferometric, radio continuum: general

1. Introduction

Source detection with radio interferometric datasets typically relies on detection of point sources in deconvolved (CLEANed) images, generated by the Fourier Transform of the underlying complex visibility dataset, observed on each base-line and at each spectral channel of the array. Detection algorithms are often applied offline with user input, and are limited by the weighting applied to the visibility data at the time of image generation; an example is SFIND, Hopkins *et al.* (2002).

Systematic study of slowly-varying transient and variable sources (slow transients, with variability time-scales >5 seconds) present an opportunity to explore the universe along a different dimension than has historically been the case (variability time-scale rather than electromagnetic spectrum), with the aim to characterize large samples of object classes, and the potential to discover new classes of objects, based on their variability properties. The next-generation of wide-field, high-sensitivity interferometric arrays have the capability to survey the sky with high cadence, yielding appropriate datasets for detecting and classifying sources varying with time-scales ranging from seconds to months. The real-time, automated detection and classification of slow transient sources from commensal observing requires statistically robust algorithms to extract information optimally, without the benefits of user input and optimal image formation. In addition, one may wish to sample the data at higher rates than the imaging time-scale, to isolate short time-scale variability. Detection of variable sources using the underlying visibility dataset is an option to interrogate the data at higher cadence, and with simpler noise properties, making optimal detection feasible. We explore visibility-based source detection in this note.

414

2. Visibility-Based Detection: the Matched Filter

For a known signal in multivariate Gaussian noise, the matched filter detector yields optimal detection performance (probability of detection maximized for a given false-positive rate). This detector can be demonstrated to be equivalent to image-based techniques. The matched filter detection performance is given by:

$$\text{SNR} = \frac{<T(x)>}{\sqrt{\text{var}[T(x)]}}, \tag{2.1}$$

where $T(x)$ is the detector test statistic, and is the output of applying the detector. Comparison of the test statistic to a threshold forms the basis of the binary signal-present/signal-absent decision. For a source of strength I_0 located at source position (l_0, m_0) embedded in white Gaussian noise, the visibility space performance has the properties:

$$
\begin{aligned}
<T(x)> &= S^\dagger S \\
&= \frac{1}{N_{\text{vis}}} \sum_u \sum_v I_0 \exp{-2\pi i(ul_0 + vm_0)} I_0 \exp{2\pi i(ul_0 + vm_0)} \\
&= I_0^2 N_{\text{vis}} \\
\text{var}[T(x)] &= <(N^T S)^2> \\
&= <S^\dagger N N^T S> \\
&= S^\dagger <NN^t> S \\
&= S^\dagger S \sigma^2/2.
\end{aligned} \tag{2.2}
$$

Therefore, the signal-to-noise ratio is given by:

$$\text{SNR}_{\text{Vis}} = I_0^2 N_{\text{vis}} / \sqrt{I_0^2 N_{\text{vis}} \sigma^2/2} = I_0 \sqrt{2N_{\text{vis}}/\sigma^2}. \tag{2.3}$$

Image-space performance relies on pixel-based detection. For a source at (l_0, m_0), a visibility is given by:

$$V(u,v) = I_0 \exp{-2\pi i(ul_0 + vm_0)} + \text{noise}. \tag{2.4}$$

The image is formed by Fourier Transform, $\vec{I} = \mathbf{H}\vec{V}$, where $\mathbf{H}$ is the Fourier operator. Therefore:

$$I(l,m) = \frac{1}{N_{\text{Vis}}} \sum_u \sum_v V(u,v) \exp{2\pi i(ul + vm)} + \mathbf{H}\mathbf{C}_N\mathbf{H}^\dagger. \tag{2.5}$$

At image position, (l_0, m_0):

$$I(l_0, m_0) = \frac{1}{N_{\text{Vis}}} \sum_u \sum_v I_0 + \mathcal{N}(0, \sigma^2/2N_{\text{vis}}). \tag{2.6}$$

Then, the signal-to-noise ratio is given by:

$$\text{SNR}_{\text{Image}} = I_0 \sqrt{2N_{\text{vis}}/\sigma^2} = \text{SNR}_{\text{Vis}}. \tag{2.7}$$

The visibility-space matched filter is therefore equivalent to pixel-based source detection in the image plane (such as that performed by SFIND), assuming that the source is located at the pixel centre, and that the visibility data have been weighted naturally (equally) in the image-formation process. These theoretical results have been presented in Trott *et al.* (2011), which also discusses source localization and optimal weighting of data for realistic noise.

3. Improved Performance

There are cases where applying a matched filter detector in visibility space will yield superior detection performance compared to traditional techniques in the image plane. Two such cases are given below.

• Image formation requires gridding the underlying visibility dataset and sampling the image plane at high spatial resolution in order to characterize adequately the instrument point spread function (synthesized beam). At the true source location and with naturally-weighted data, the amplitude of the image-space pixel contains all the information about the source. For a source not co-located with a pixel, the flux is spread across pixels and pixel-based detection no longer contains all the information. Visibility-based detection does not suffer from such discretization errors;

• Commensal observing modes for detecting and classifying slow transients will be routine for next-generation instruments. In that mode, the image-formation process will be designed to be optimal for the science case being studied. The images may not be weighted for optimal point-source sensitivity. In addition, if the underlying noise in the visibility dataset is not white Gaussian (i.e., it is coloured or correlated), failure to include that in the image-formation process will lead to degraded performance. Operating directly on the visibility data allows optimal weighting of the data in the source detector to be achieved.

4. Summary

The detection of previously unknown sources, whether static, variable or transient, is one of the key goals of future wide-field, high-sensitivity, large interferometric radio instruments. To realise the potential to explore the unknown parameter space of variable radio objects and to discover new classes of such objects, we require sophisticated source detection, variability detection and source-localization techniques. The large data rates and advanced scientific demands of these instruments will require those detections and classifications to be performed in real time, automatically, and often in a commensal mode of observing. Those constraints demand novel thinking about how to design optimal pipelines to extract as much information before data are destroyed. Working in visibility space allows high-time-resolution access to the raw data in a form with maximal information content and simple noise statistics. In this paper we have explored the theoretical possibility of achieving optimal source detection with visibility data, and demonstrate its equivalence to image-based techniques, for the case of simple noise properties. Details of the source-localization algorithm and visibility-space detection performance will appear in a forthcoming publication.

References

Hopkins, A. M., *et al.* 2002, *AJ*, 123, 1086
Trott, C. M., Wayth, R. B., Macquart, J.-P., & Tingay, S. J. 2011, *MNRAS*, 731, 81