Draft version, To appear in URSI GASS 2014, (C) IEEE 2014

ExCon: Making Billion Pixel Radio Interferometric Images

Sarod Yatawatta,
ASTRON, The Netherlands Institute for Radio Astronomy,
The Netherlands.
yatawatta@astron.nl

Abstract

Imaging a wide field of view with high enough resolution is needed for calibration and achieving a high dynamic range, especially with low frequency interferometers. This paper describes a new imager (ExCon) that is able to make widefield images that are far larger in size compared to images made by existing imaging software. In the process of implementing this new imager, we have also introduced new algorithms in image weighting and convolution kernel calculation. We also give details about these algorithms. Deconvolution algorithms will be implemented as future work.

I. Introduction

Widefield imaging is an imperative in low frequency radio astronomy. Getting an accurate model of a large fraction of the sky is essential for making high dynamic range images [1]. With long baselines, these images require a high resolution. Wide field of view combined with high resolution implies making images that are large in size. In fact, the number of pixels can easily run up to one billion. Images made using multiple facets are not a viable solution because of edge effects. Sources at the edge of facets get distorted and it is hard to extract good source models for them. Therefore, we need to create single facet images. For instance, for a LOFAR high band antenna observation at 150 MHz, the field of view is about 20 degrees and at $2''$ resolution, the image size is 36 000$\times$36 000, which runs up to about 1.3 billion pixels.

Existing imaging software runs out of memory while making such images. Therefore, in this paper, we present a new imager (ExCon) especially created to make such images using a modest amount of memory. In the following, we describe various stages of imaging (weighting, convolutional gridding, and Fourier transforming) and how the new imager is different from existing imaging software. In addition, we also introduce a new weighting scheme that is suitable for minimizing PSF (point spread function) sidelobe variation while maximizing the signal to noise ratio.

II. Image weighting

Interferometric data are taken in the Fourier (uv) space and irregularly sampled due to the array configuration. Let $s(x)$ be the signal in the image space (coordinate x) and its Fourier transform $S(k)$ in the Fourier space (coordinate k). The values of k are irregular and depends on the uv coverage of the instrument. Moreover, the values of k scale with frequency of the observation. Throught this paper we consider a 1 dimensional analog, which is easily extended to higher dimensions. The weights $W(k)$ are applied to the data to get $S(k)W(k)$ which is used in imaging.

An overview of existing weighting schemes can be found in [2] and [3]. Commonly used 'natural' weights make $W(k)=1$ while 'uniform' weights calculate $W(k)$ depending on the number of data points within a gridded cell. In contrast, there is a wide variety of iterative weighting techniques [4],[5],[6] that are used in magnetic resonance imaging. We adopt these techniques to radio interferometric imaging. Let the convolution kernel in k space be given as $C(k)$, then after convolution, we have $(S(k)W(k))\otimes C(k)$ which is sampled on a regular grid to perform the fast Fourier transform (FFT). In image space, after convolution, we have $(s(x)\otimes w(x))\,c(x)$, although we cannot calculate this in practice. The motivation for [4] to derive an iterative weighting scheme is as follows. Assume we need to closely approximate a desired function $g(x)$ in image space,

$$(s(x)\otimes w(x))\,c(x)\approx g(x). \tag{1}$$

Note that $S(k)$ is not necessarily the observed data but a positive, real valued function, for example, representing the sensitivity of each datapoint. Also $W(k)$ is positive and real valued, because it is the weight. We convolve both sides of (1) by the image space equivalent of $W(k)$, i.e. $w(x)$ to get

$$w(x)\otimes((s(x)\otimes w(x))\,c(x))\approx w(x)\otimes g(x) \tag{2}$$

where $c(x)$ is the image space equivalent of $C(k)$. Taking Fourier transform of both sides of (2), we get

$$W(k)\,((S(k)W(k))\otimes C(k))\approx W(k)G(k) \tag{3}$$

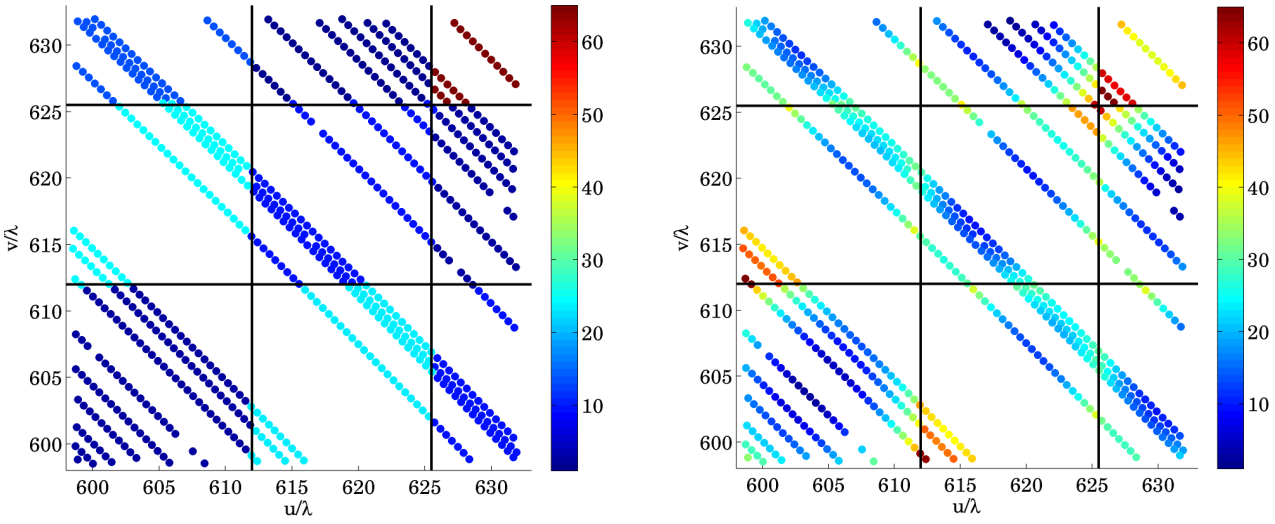
where $G(k)$ is the Fourier transform of $g(x)$, and assuming $((S(k)W(k)) \otimes C(k))$ is finite, we divide (3) by this term to get

$$W(k) \approx \frac{W(k)G(k)}{((S(k)W(k)) \otimes C(k))}. \tag{4}$$

We can use (4) to find $W(k)$ such that, when applied to the data, we can closely approximate $g(x)$. Iteratively, this can be expressed as

$$W^{i+1}(k) \leftarrow \frac{W^i(k)G(k)}{((S(k)W^i(k)) \otimes C(k))} \tag{5}$$

where $W^i(k)$ is the weight at the i-th iteration. We calculate the right hand side of (5) with the weights at the i-th iteration and update the weights by taking the ratio. In other words, given $S(k)$ and $G(k)$, we can iteratively find $W(k)$ to satisfy (1). For instance, when $S(k) = 1$ and $g(x) = \delta(x)$ which is the Dirac delta function, we get 'uniform' weights for $W(k)$. In Fig. 1, we show weights on 3×3 grid cells in the Fourier space, where traditional 'uniform' weights and weights obtained by (5) are applied.



(a) (b)

Fig. 1. Image weights (log-scale weights normalized to match linear color scale) on an area covering 3×3 grid cells. (a) Uniform weights, where points within each grid cell are assigned equal weights. (b) Iterative weights obtained by (5), after 3 iterations. Within a grid cell, the weights are adjusted according to the density of sampling points.

We see that in Fig. 1 (a), equal weights are assigned to points within a grid cell. However, in Fig. 1 (b), the weights are adjusted to match the density of points. Both weighting schemes give almost identical images but the weights in Fig. 1 (b) increase the noise because points within a grid cell are unevenly weighted.

Taking the idea of [4] a bit further, we consider the applicability of (5) for deep imaging experiments [7],[8] and overcoming some problems due to PSF variation. For such experiments, it is important to have the same PSF over a wide frequency range, and over many observations taken on different days. Moreover, due to the weakness of the sought after signal, it is also important to maximize the signal to noise ratio. For instance even when using conventional 'uniform' weights, there is an increase in noise and significant variation in the PSF due to the scaling of k coordinate.

There is another aspect that needs to be considered. Let the weak (and large scale) signal we are after be $f(x)$, and the PSF be $g(x)$. For Lebesgue measurable $f(x)$ and $g(x)$, applying Young's inequality [9] we get

$$\|f \otimes g\|_r \leq \|f\|_p \|g\|_q, \quad \frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1, \ 1 \leq p, q, r \leq \infty \tag{6}$$

where $\|f\|_p$ is the p-norm,

$$\|f\|_p = \left(\int_\Omega |f(x)|^p dx \right)^{\frac{1}{p}}. \tag{7}$$

Considering that the image we make is $f \otimes g$ (deconvolution is not possible because $f(x)$ is way below the noise level), we can enhance the properties of $f(x)$ by properly selecting $g(x)$.

Due to both aforementioned reasons, the importance of properly selecting $g(x)$ becomes apparent and the scheme proposed in (5) allows us to do this. For instance, we can select $g(x)$ such that $G(k)$ is the histogram of the sampling points in Fourier space over all frequencies and all different days when data was taken. With this fixed $g(x)$, we calculate unique $W(k)$ for each dataset and each frequency. The only requirement for $g(x)$ is that it should be Fourier transformable, i.e. has finite energy.

For completeness, we give the remaining steps in making an image. First, the convolved data is sampled on a regular grid, $\left(\left(\tilde{S}(k)W(k)\right) \otimes C(k)\right) \text{Ш}(k/B)$ where $\text{Ш}(k/B)$ is the Dirac comb with uv cell size B. Note that $\tilde{S}(k)$ (and $\tilde{s}(x)$) is the actual observed data. After taking the FFT, we have $\left(\left(\tilde{s}(x) \otimes w(x)\right) c(x)\right) \otimes \text{Ш}(Bx)$ with aliased copies of the image spaced at $1/B$. Finally, after apodization correction, we get $\left(\left(\tilde{s}(x) \otimes w(x)\right) c(x)\right) \otimes \text{Ш}(Bx)/c(x)$ as the final image. If B is small enough, (and $1/B$ large enough), we can ignore the aliased copies to get $\left(\tilde{s}(x) \otimes w(x)\right)$ as the final image. This is actually the true signal $\tilde{s}(x)$ convolved with the PSF, $w(x)$.

III. Convolutional kernel calculation

Starting with [10],[11], prolate spheroidal wave functions (PSWF) have been used as the convolution kernel $C(k)$ in all imaging software. This, coupled with w-projection [12] is used due to the 3 dimensional Fourier space sampling. In [11], PSWF are calculated using a rational polynomial approximation. However, in ExCon, we use [13] to exactly calculate the PSWF using an eigenvalue decomposition. This way, we are able to vary the bandwidth and support length of PSWF depending on the requirement. Moreover, we use both w-projection and w-snapshots in ExCon.

IV. Fourier transform

Most demanding memory requirement for imaging is for taking the FFT. However, we use memory-mapped input/output to store the gridded data and the output of the FFT on disk. We use FFTW library [14] to evaluate the FFT. This library does not make any distinction between primary memory and disk as input/output while performing the FFT. Therefore, we are not limited by memory for the size of the FFT. With access times to disk ever decreasing (such as solid state drives), we do not see a major performance difference for smaller images and for larger images, disk storage still works while primary memory will be exhausted.

V. Results

We give an example of iterative weighting to match $g(x)$ which is made to be a Gaussian with full width half maximum (FWHM) of $5'$ and $10'$ in Fig. 2. We have compared this result with conventional Gaussian tapering with a Gaussian of same size.

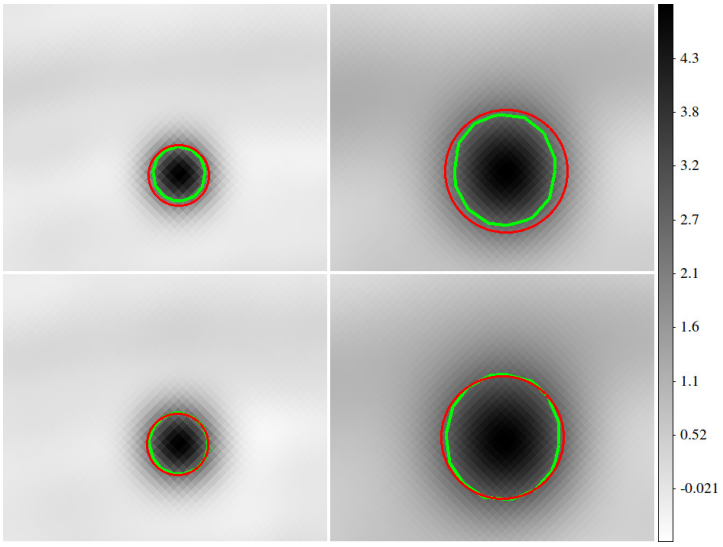


Fig. 2. Iterative weighting compared with tapering. The top row shows images of a point source after tapering. The bottom row shows the same image after application of iterative weights to match $g(x)$. For the iterative weighting, we started with $W(k) = 1$ and performed 3 iterations. The left column shows images for a Gaussian of $5'$ FWHM while the right column $10'$ FWHM. The red circles correspond to the desired FWHM of the Gaussian while the green contour indicates the half maximum level in the image. We see that the actual contour matches the desired value better in the bottom row, indicating that iterative weighting does a better job in finding the weights as opposed to tapering.

The images show a point source, after 'uniform' weighting and tapering with a Gaussian of FWHM of $5'$ and $10'$. We also show the same point source, by iteratively selecting the weights as in (5) to make $g(x)$ equal to the same Gaussian with

FWHM of $5'$ and $10'$. We see from the images in Fig. 2 that iterative weighting gives better accuracy in attaining the desired shape for $g(x)$.

The largest image that we have made by using ExCon is of size 36 000×36 000 pixels, at $2''$ resolution, as shown in Fig. 3. The size of the FFT, taking into account the padding, is 45 000×45 000 pixels. This was done on a computer with only 12GB of memory and 48GB temporary disk storage was used. The total time required was about 24 hours. The bottleneck is the gridding time, and in order to reduce this, for w-projection, we only used 128 pixels for the convolution kernel support. The secondary bottleneck is the disk input/output time, and we are investigating various orderings of data to optimize this.

VI. Conclusions

We have presented a new imager, ExCon, and introduced a new image weighting scheme. As future work, we consider using distributed memory (across a network of computers) for performing the FFT. The weight calculation and gridding operations are being accelerated using GPUs. Moreover, we are also considering different data ordering schemes to optimize performance. Deconvolution algorithms will also be incorporated to the imager at a later stage.

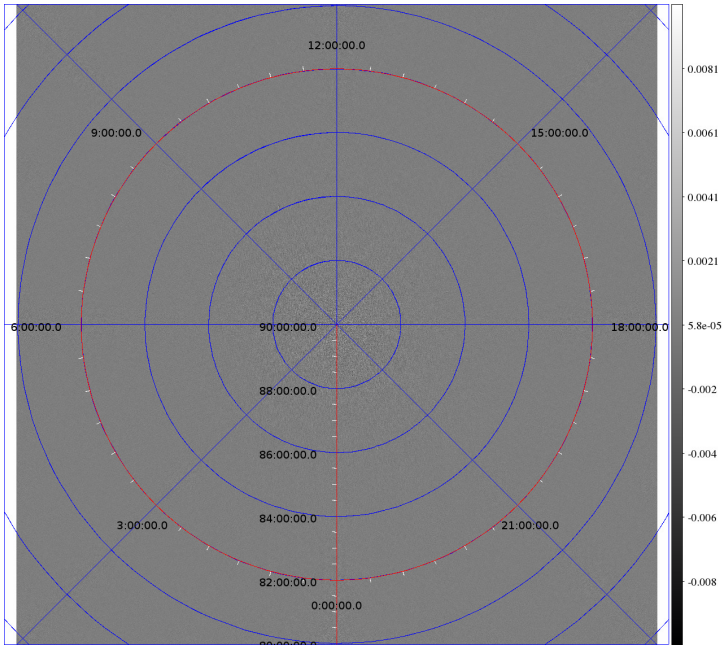


Fig. 3. An image made with ExCon of a simulated dataset. The field of view is 20×20 degrees, and at $2''$ pixel size, the image size is 36 000×36 000 (or 1.3 billion pixels). It is hard to see any detail due to the low resolution on this paper.

References

[1] J. Bregman, "System design and widefield imaging aspects of synthesis arrays with phased array stations," *PhD Thesis, Univ. Groningen*, Dec. 2012.

[2] D.S. Briggs, "High fidelity deconvolution of moderately resolved sources," *PhD Thesis, New Mexico Institude of Mining and Technology*, 1995.

[3] F. Boone, "Weighting interferometric data for direct imaging," *Experimental Astronomy*, vol. 36, no. 1-2, pp. 77–104, 2013.

[4] J.G. Pipe and P. Menon, "Sampling density compensation in MRI: Rationale and an iterative numerical solution," *Magnetic Resonance in Medicine*, vol. 41, pp. 179–186, Dec. 1999.

[5] F. Wajer, R. Lethmate, R. de Jong, L. Martinez, D. Graveron-Demily, M. Fuderer, and D. van Ormondt, "MR image reconstruction from sparse and corrupted k-space data," in *ProRISC, IEEE Benelux, Mierlo*, 1999, pp. 577–584.

[6] K.O. Johnson and J.G. Pipe, "Convolution kernel design and efficient algorithm for sampling density correction," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 439–447, 2009.

[7] S. Zaroubi, A. G. de Bruyn, G. Harker, R. M. Thomas, P. Labropolous, V. Jelić, L. V. E. Koopmans, M. A. Brentjens, G. Bernardi, B. Ciardi, S. Daiboo, S. Kazemi, O. Martinez-Rubi, G. Mellema, A. R. Offringa, V. N. Pandey, J. Schaye, V. Veligatla, H. Vedantham, and S. Yatawatta, "Imaging neutral hydrogen on large scales during the Epoch of Reionization with LOFAR," *Mon. Not. R. Astron. Soc.*, vol. 425, pp. 2964–2973, Oct. 2012.

[8] B. J. Hazelton, M. F. Morales, and I. S. Sullivan, "The Fundamental Multi-baseline Mode-mixing Foreground in 21 cm Epoch of Reionization Observations," *The Astrophysical Journal*, vol. 770, pp. 156, June 2013.

[9] J.F. Fournier, "Sharpness in Young's inequality for convolution," *Pacific J. Math.*, vol. 72, no. 2, pp. 383–397, 1977.

[10] W. N. Brouw, "Aperture synthesis," *in Methods in Computational Physics*, vol. 14, pp. 131–175, 1975.

[11] F.R. Schwab, "Optimal gridding," *VLA Scientific Memorandum 132*, pp. 1–18, 1980.

[12] T.J. Cornwell, K. Golap, and S. Bhatnagar, "The noncoplanar baselines effect in radio interferometry: The w-projection algorithm," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 5, pp. 647–657, Oct 2008.

[13] J. Xiao, V. Rokhlin, and N. Yarvin, "Prolate spheroidal wavefunctions, quadrature and interpolation," *Inverse Problems*, vol. 17, pp. 805–838, Dec. 2001.

[14] Matteo Frigo and Steven G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005, Special issue on "Program Generation, Optimization, and Platform Adaptation".