Astronomical Data Analysis Software and Systems X
ASP Conference Series, Vol. 238, 2001
F. R. Harnden Jr., F. A. Primini, and H. E. Payne, eds.

Parallelization of Widefield Imaging in AIPS++

K. Golap, A. Kemball, T. Cornwell, W. Young

National Radio Astronomy Observatory

Abstract. At low frequencies, large synthesis arrays in Radio Astronomy, such as the Very Large Array (VLA), effectively require that a 3-D Fourier transform be used in imaging, rather than the conventional 2-D transform. Given the large data volumes associated with observations of this type, this ensures that these problems are amongst the most computationally demanding in radio astronomy. Typical image sizes are of the order of a few million pixels.

The wide-field imaging problem can be made more tractable by using parallelization. In this paper, we discuss the general wide-field imaging algorithm used in AIPS++, and the techniques used for its parallelization.

1. Overview of the Wide-Field Imaging Problem

A problem occurs when imaging large fields of view with relatively long baselines and non-coplanar arrays. Imaging using synthesis arrays involves inverting the 3-D integral

$$V(u,v,w) = \int I(l,m) \exp j2\pi(ul + vm + wn) \frac{dldm}{\sqrt{1-l^2-m^2}}$$

to obtain the brightness distribution $I(l,m)$ on the sky, from a measured set of visibilities, $V(u,v,w)$, in the uv-plane. In most practical cases the non-coplanar term w can be neglected and the inversion is a direct 2-D Fourier transform. However, for wide-field imaging if the w term is not taken into account there is usually a substantial loss of dynamic range, and it is also impossible to faithfully image regions far from the field center.

2. The Widefield Imaging Algorithm Used in AIPS++

Several algorithms exist to solve the full 3-D problem listed above (Cornwell & Perley 1992). In AIPS++ a multi-faceted transform approach has been chosen for its efficiency. This covers the image plane by a series of facets, in each of which a 2-D transform holds.

We can decompose the visibilities into a summation of re-phased faceted visibilities:

$$V(u,v,w) = \sum_k V_k(u,v) \frac{\exp j2\pi(ul_k + vm_k + w\sqrt{1-l_k^2-m_k^2})}{\sqrt{1-l_k^2-m_k^2}}$$

where :

$$V_k(u,v) = \int I_k(l - l_k, m - m_k) \exp j2\pi(u(l - l_k) + v(m - m_k))dldm$$

The iterative multi-stage algorithm implemented in AIPS++ proceeds as follows:

- Calculate residual images for all facets (using 2-D transforms).
- Partially deconvolve individual facets and update the image model for each facet.
- Reconcile different facets by subtracting the model visibility for all facet models from the visibility data.
- Recalculate residual images and repeat. In the process of making residual images, a uv-plane coordinate system is chosen so that the final image from all facets is projected on a common tangent plane (Sault et al. 1996).

2.1. An Example

Wide-field imaging is computationally expensive. The image in Figure 1 was made using the AIPS++ widefield algorithm with 225 facets. The data are a VLA observation at 74 MHz in the B and C configurations. This image took close to 20 days to process on a desktop workstation (SGI octane). A similar observation in the A-array of the VLA would require some ten times more computer resources to process. Along with other overheads, like better deconvolution algorithms for larger baselines, we are facing computation of 200 to 300 days on a typical desktop. This problem strongly justifies the need for parallelization of this algorithm. The problem will be more pronounced with future arrays such as the Expanded VLA (EVLA).

3. The Parallelization Effort and Progress

For the first level of parallelization we are aiming at parallelizing the nearly embarrassingly parallel sections of the widefield algorithms. There are three distinct sections which we have identified in the widefield algorithms which fall under this category:

- The point spread function (PSF) formation. The PSF for each facet is needed in deconvolution. These can be estimated totally independently of each other, requiring only the uv-coverage seen from each facet.
- The model visibility estimation from the source model components. As the visibilities from different sources (or different facets) are additive, they can be estimated independently for each facet model and cumulatively added into the final model visibility. This has parallel I/O implications.
- The residual image estimation. In calculating the residual image the residual visibility re-projection for the different facets can be estimated independently.

3.1. Progress Made So Far

We have made progress in parallel I/O development and evaluation of different access methods for the visibility data. This includes measuring the efficacy of

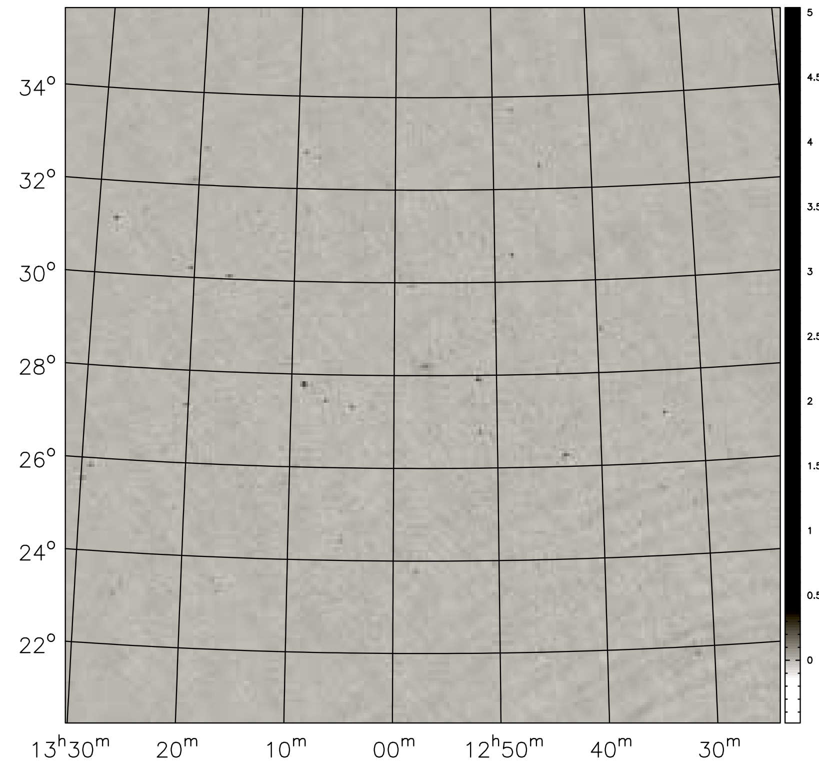

Figure 1. Image of Coma cluster at 74 MHz reduced in AIPS++ using 225 facets

parallelization with multiple processes accessing the same visibility data. We have verified that parallelization of the function to form the PSFs speeds up almost linearly for a few processors, and have parallelized the function to predict the model visibilities.

4. Ongoing and Future Work

The parallel 3-D imaging approach is close to full operational use. Areas of ongoing work include: migration to larger machines or clusters and fully measuring the speed up of each algorithm computed, parallelizing the residual image formation for each facet. further work in Parallel I/O using MPI-2, and investigation of statement level parallelization using OPEN-MP.

References

Cornwell, T. J. & Perley, R. A. 1992 Astron. & Astrophys, 261, 353

Sault, R., Staveley-Smith, L., & Brouw, W. N. 1996 Astron. & Astrophys. Suppl., 120, 375