Summary of Proposed Research Program

High-Time Resolution GPU Imager for Low-Frequency Radio Telescopes

Applicant:
Gayatri Aniruddha

Supervisor:
Dr. Marcin Sokolowski

Curtin Institute for Radio Astronomy
Masters of Philosophy (MPhil)

Application for Candidacy

Abstract

Fast Radio Bursts (FRBs) are very bright (even tens of Jansky) milli-second radio pulses originating even from a very distant Universe. The search for FRBs is similar to that of single pulse search from pulsar observations through beam-forming. This method has a high associated computational cost and hence, in my research project, I am exploring alternate image-based approaches, specifically for low-frequency FRB searches. Furthermore, the reduced signal to noise ratio at low frequencies due to long dispersive delays and sky background temperature makes FRB detections at low frequencies even more demanding. Existing imagers such as CASA, MIRIAD and WSCLEAN have substantial I/O heavy operations, while high-time resolution images are required for FRB searches. Furthermore, there is no imaging software, operating entirely or predominantly on GPUs to look for FRBs. Thus, the main aim of my Masters by Research is to develop an alternative GPU imager which will be a part of the processing pipeline searching for FRBs. I will continue working on the imaging software to test and deploy it on SKA-Low stations and the Murchison Widefield Array (MWA) and further optimise its processing efficiency by implementing it on GPUs to capitalise on parallel programming. Finally, this imager will be used to image(in 10 or 100 ms time resolution) small amount (atleast 10 min, up to 60min) of SKA-Low station or MWA data to demonstrate the potential of the imager for high-time resolution science.

International Centre for Radio Astronomy Research

Curtin University

Contents

1 Objectives

Since the first FRB discovered in 2007 (Lorimer et al., 2007), the FRB field has swiftly progressed to a state where detections are now made in real-time by several telescopes worldwide. (Petroff et al., 2022). Earlier, the detections were made at 1-2 GHz and in spite of substantial searches over the last years, only few FRBs have been clearly detected below 350 MHz and so far, no FRB has been detected with the MWA (Tingay et al., 2013) or with the SKA-Low stations (Wayth et al., 2022; Macario et al., 2022). Existing imagers such as the Common Astronomy Software Applications (CASA, CASA Team et al., 2022), Multi-channel Image Reconstruction, Image Analysis, and Display (MIRIAD, Sault et al., 2011) and WSCLEAN (Offringa et al., 2014) are not best suited to look for FRBs as high-time resolution(100 ms, down to 10 ms) images are needed for look for FRBs. Additionally, though there is an image-domain-gridding mode (IDG, van der Tol et al., 2018) in WSCLEAN, which can operate in either a CPU mode/GPU mode or in an hybrid mode, the 'pure' GPU option is not recommended as it does not contain all the features of the hybrid mode. Thus, no existing imagers operate completely on GPUs which if implemented, could potentially fasten the process of imaging. Therefore, this implementation of imagers completely on GPUs, for both SKA Low stations and the MWA is going to be the main focus of my Masters by Research project.

2 Background

The main goal of my research is to develop a GPU imager applicable to FRB searches. Hence, in this section I have provided a brief overview of the FRB field in general, FRB surveys at low frequencies and the computational cost of data processing and searches.

2.1 Fast Radio Bursts

FRBs are short duration (milli seconds) radio pulses with high flux densities, originating from a very distant Universe. The first FRB also called the Lorimer burst was discovered in 2007, while processing archival pulsar data using the Parkes radio telescope (Lorimer et al., 2007). This burst had a dispersion measure (DM) of 375 pc cm^{-3} and a flux density of 30 Jansky. The DM is defined as the integral of the electron number density along the line of sight, and is directly related to the distance of the FRB. When FRBs travel from their source, they undergo dispersion, a mechanism where the group velocity of the wave varies according to its frequency which results in signals with a higher frequency to arrive earlier and signals with a lower frequency to arrive later, which leads to a delay in their arrival times (Δt) proportional to ν^{-2}.

2.2 Fast Radio Bursts at frequencies below 400 MHz

The current total number of FRB detections made as of January 2023 stand at 639, which include a total 50 repeaters (Petroff & Chatterjee, 2023). Most FRBs appear to

be one-off (or apparently non-repeating) events i.e they have been detected only once at a given position. Some FRB are known to repeat, where bursts are seen from the FRB source at irregular intervals. FRB 121102 (Spitler et al., 2016), discovered by the Arecibo Radio Telescope (Cordes et al., 2006) in Puerto Rico was the first repeating FRB to be discovered. FRB 20180916B, a well-known and the most precisely localised repeating FRB was discovered by the Canadian Hydrogen Intensity Mapping Experiment FRB backend (CHIME/FRB, CHIME/FRB Collaboration et al., 2019). Although many FRBs have been seen to repeat at irregular intervals, only this FRB has been detected to repeat in a regular way where it seems to repeat every 16.35 days. This detection of periodicity from FRB 20180916B was a milestone in the history of FRB detections because more telescopes around the world were utilised in a productive manner in this pointed direction to detect bursts from this FRB even down to ∼100 MHz frequencies (Pleunis et al., 2021).

2.3 Previous FRB searches using SKA-Low stations and MWA

The MWA, located at the Murchison Radio-astronomy Observatory (MRO) in Western Australia is a low frequency radio telescope with a wide frequency operating range of 70-300 MHz and a huge field-of-view (hundreds of square degrees), making it very well suitable for a large-scale survey of the sky for low-frequency FRBs. Though the previous searches for FRBs using MWA were successful in terms of establishing upper limits on the sensitivity of the searches, till date no FRB has been discovered by the MWA or even by the SKA-Low stations.

2.4 Current methods for FRB searches

FRB searches are similar to that of single pulse searches from pulsar observations, where data are prepared for searches through beamforming. In beamforming, voltages recorded by the telescope are combined into a single tied-array beam by performing a coherent addition of voltages over all the antenna elements by applying suitable delay and phase corrections. Therefore, the total computational cost associated to produce an image in a single direction is directly proportional to the number of antennas present in the telescope and desired image size.

3 Significance

The associated computational cost to produce image results from the main three procedures of correlation ($0.5N_a(N_a - 1)$), gridding ($N_{vis}N_{kern}^2$) and Fourier transform ($N_{px}^2 log_2(N_{px}^2)$), explained further in Section 4.1, is typically lesser than the cost of forming multiple tied-array beams ($N_a N_{pixels}^2$), where N_a refers to the number of antennas, N_{px} refers to the number of pixels in the desired image size, N_{vis} are the number of visibilities that need to be gridded and N_{kernel} is the convolution kernel used for gridding. Additionally, though existing imagers such as CASA, MIRIAD and WSCLEAN are used

for generating images for FRB searches, they have substantial I/O operations. Furthermore, other than the IDG option in WSCLEAN, these imagers do not have the option of GPU implementations. Thus, a GPU-based imager that generates high-time resolution images is needed, as it can process data faster to generate images and potentially improve computational efficiency and yield of FRB searches at low frequencies.

4 Proposed Research Method

In this research project, I am looking at image-based approaches, implemented using GPUs to prepare data for FRB searches, suitable for both SKA Low stations and the MWA. The initial phase of this research is to profile the current version of the CPU imager which generates images from one time and frequency channel to identify the most time consuming sections to further implement them on GPUs and benchmark its performance. In the next step, a GPU imager will be developed for one frequency channel and observing time, for EDA2 data and then further expanded to incorporate multiple frequency channels and observing times by processing data in parallel. In this section, I have started by giving an overview of the visibility function, followed by description of radio astronomy imaging. At the end, I have also described the work completed up until now and my plans for my first proposed paper.

4.1 Imaging in Radio Astronomy

Imaging of radio data is the process of creating an image of the sky at radio frequencies, starting from voltages recorded by multiple antennas in a radio telescope. An image is created for a given frequency and a given point in time. The voltages from antennas are correlated to obtain 'visibilities', a complex quantity expressed as a function of baseline coordinates, measured in wavelengths. It is mathematically expressed as:

$$V(u,v) = \iint T(l,m) e^{-i2\pi(ul+vm)}\, dl\, dm, \tag{1}$$

where *V(u,v)* is the visibility function and *T(l,m)* is the sky brightness distribution, yielding the image, *u,v* are East-West, North-South spatial frequencies in wavelengths, *l,m* are East-West, North-South angles in the tangent plane in radians.(Thompson et al., 1991). In order to get what is called a 'dirty image' of the sky, in theory, Fourier transform is applied on the visibilities to transform data from the uv-plane to the image plane. However, in practice, the fast Fourier Transform (FFT) algorithm is used to make it computationally efficient. FFT however requires data to be to be placed on a regularly spaced grid. Therefore, a method called 'gridding' is implemented where all the visibilities are placed on a regularly spaced grid, the dimensions of which are dependent on the size and dimensions of the requested output sky image. After visibilities are gridded on the uv-plane, FFT is applied to transform data from the uv-plane to the image plane, which results in the sky image.

4.2 Work to date

Till date, I have been working on the GPU imager proposed for Paper I and the following sections highlight the different tasks I have completed towards this paper.

1. The current version of the imager is implemented completely in `C++` and generates all-sky images using EDA2 data for one frequency channel and observing time. As a part of the first stage of my project, I implemented the 'gridding' procedure through a `CUDA` kernel using atomicAdd() (Atomic Operations). I then implemented the 'imaging' procedure using `cuFFT` (cuFFT API Reference). The resulting GPU image is shown in Figure 3 in the Appendix section. The image was verified by comparing it with the image generated with the original CPU version and it was very similar to it. However, small fractional differences of around 0.25 and -0.5 were identified in the real and imaginary gridded visibilities respectively due to to different (pseudo-random) order of addition operations in GPU version of gridding. This resulted in small differences of 0.0011 and 0.0009 in the final real and imaginary dirty images respectively. It is currently investigated if these differences are acceptable.

2. In order to benchmark the performance of implementing gridding and imaging in GPUs, I profiled both CPU and GPU imager using `clock()` functions for different image sizes as shown in the figure below. Figure 1 captures the imager performance with and without GPU implementation, and there is a noticeable improvement in the execution time in the GPU Imager for higher image sizes.

3. Additionally, I also profiled this GPU imager using `nvprof` for different image sizes (180×180, 1024×1024 and 8182×8182) and different number of executions (1, 2 and 10). In all the cases, majority of the GPU time was spent in memory transfers between the Host(CPU) and the Device(GPU) and in running the `CUDA` kernel. The detailed results are captured in Figure 2 in the Appendix section. These memory transfers will further increase with an increase in the number of frequency channels and observing times. Thus, the next step in optimisation would be to try minimising these memory transfers between host and device as much as possible by copying larger blocks of data (corresponding to multiple time steps and/or frequency channels) in GPU memory, before copying the results back to the CPU memory.

4.3 Paper I: High-time resolution GPU imager for Low-frequency Radio Telescopes.

The focus of this paper will be the description of the GPU imager for SKA Low Stations and MWA for the detection of FRBs. This work has already started by implementing the sections of gridding and imaging on GPUs, for one frequency channel and observing time, using EDA2 data. This will be further developed to incorporate multiple frequency channels and observing times. The imager will also be tested and validated on simulated

data in order to demonstrate its correctness, the results of which will be described in this paper. We will compare the resulting images to corresponding images produced by other radio astronomy packages such as CASA, MIRIAD or WSCLEAN. These comparisons may be extended to performance comparisons but it will depend if we can compare relatively basic functionality (dirty image) to corresponding functionality of existing imagers in a "fair way". Furthermore, this paper will also capture the performance reports of the GPU imager, profiled on both the Topaz and Setonix Supercomputers at PAWSEY. This imager will be further used to image (in 10 or 100 ms time resolution) small amount (at least 10 min) of EDA2 or MWA data. We will select data, which will demonstrate the potential of the imager for high-time resolution science, for example by demonstrating detections of bright pulses from a known pulsar or showing high-time resolution images of radio-frequency interference (RFI) caused by natural (e.g. lightning) or human sources (satellites or other transmitters).

5 Ethical Issues

This research does not involve human or animal subjects. Therefore, there are no ethical issues applicable for this research.

6 Resources, Facilities and Data Storage

I will be using data from SKA Low Stations and MWA for my Masters by Research. Topaz and Setonix Supercomputers located at Pawsey will be used for code development and testing. Additionally, GitLab and will used for version control. Furthermore, my desktop at CIRA/ICRAR and One Drive will be used for keeping a back-up of the code, reports, papers and other relevant documentation.

7 Timeline

		Months since 10 December 2021							
Activity		3	6	9	12	15	18	21	24
Literature review									
Learning software packages									
Papers	I - High Time Resolution GPU Imager for SKA-Low stations								
Thesis	Chapter 1 - Background								
	Chapter 2 - Methods								
	Chapter 3 - Paper 1								
	Chapter 4 - Discussions and Conclusion								

Table 1: Timeline

8 Appendix

Image size	180 x 180	256 x 256	512 x 512	1024 x 1024	2048 x 2048	4096 x 4096	8182 x 8182
CPU Total Execution Time	0.02	0.07	0.11	0.36	1.82	8.26	43.48
GPU Total Execution Time	0.42	0.41	0.47	0.58	1.42	4.68	13.2
CPU/GPU	0.05	0.17	0.23	0.62	1.28	1.76	3.29

Figure 1: A comparison of the execution times of CPU and GPU implementations of the imager, tested on EDA2 data.

GPU Activities / Image Size	180 x 180	1024 x 1024	8182 x 8182
CUDA memcpy HtoD	95.966	96.159	139.77
CUDA memcpy DtoH	13.248	4.1966	347.46
gridding_imaging_cuda	19.616	19.519	20.639

Figure 2: The benchmarks in micro-seconds of different GPU activities obtained using `nvprof` for various image sizes and for a single execution. The first two GPU activities, namely CUDA memcpy HtoD and CUDA memcpy DtoH refer to the memory transfers between CPU to GPU and GPU to CPU respectively. The third GPU activity, gridding_imaging_cuda is the function which performs the procedures of gridding and imaging in CUDA.

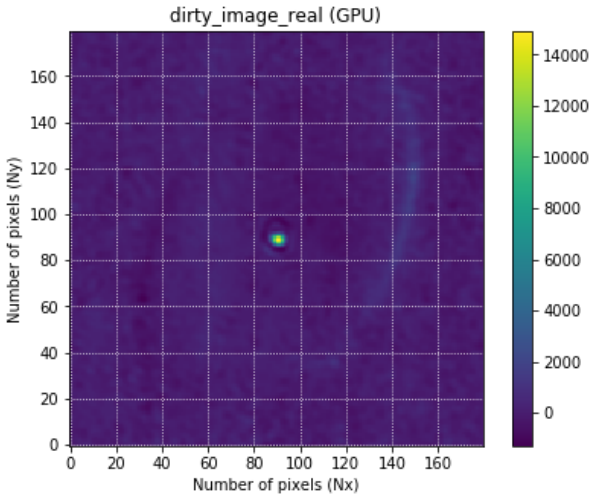

Figure 3: An all-sky image of size, 180 × 180, produced by the current GPU Imager from EDA data. The centre object is the sun and the stripe on the left is the galactic plane.

References

CASA Team, Bean, B., Bhatnagar, S., et al. 2022, PASP, 134, 114501, doi: 10.1088/1538-3873/ac964210.48550/arXiv.2210.02276

CHIME/FRB Collaboration, Andersen, B. C., Bandura, K., et al. 2019, ApJ, 885, L24, doi: 10.3847/2041-8213/ab4a80

Cordes, J. M., Freire, P. C. C., Lorimer, D. R., et al. 2006, ApJ, 637, 446, doi: 10.1086/498335

Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Science, 318, 777, doi: 10.1126/science.1147532

Macario, G., Pupillo, G., Bernardi, G., et al. 2022, Journal of Astronomical Telescopes, Instruments, and Systems, 8, 011014, doi: 10.1117/1.JATIS.8.1.01101410.48550/arXiv.2109.11983

Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, MNRAS, 444, 606, doi: 10.1093/mnras/stu136810.48550/arXiv.1407.1943

Petroff, E., & Chatterjee, S. 2023, FRB Newsletter Volume 04, Issue 01 — January 2023. https://ecommons.cornell.edu/bitstream/handle/1813/112766/v4-01_January2023.html?sequence=2&isAllowed=y

Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2022, , 30, 2, doi: 10.1007/s00159-022-00139-w

Pleunis, Z., Michilli, D., Bassa, C. G., et al. 2021, ApJ, 911, L3, doi: 10.3847/2041-8213/abec72

Sault, R. J., Teuben, P., & Wright, M. C. H. 2011, MIRIAD: Multi-channel Image Reconstruction, Image Analysis, and Display, Astrophysics Source Code Library, record ascl:1106.007. http://ascl.net/1106.007

Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Nature, 531, 202, doi: 10.1038/nature17168

Thompson, A., Moran, J., & Swenson, Jr, G. 1991, Interferometry and Synthesis in Radio Astronomy, Vol. -1, doi: 10.1007/978-3-319-44431-4

Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, e007, doi: 10.1017/pasa.2012.007

van der Tol, S., Veenboer, B., & Offringa, A. R. 2018, A&A, 616, A27, doi: 10.1051/0004-6361/201832858

Wayth, R., Sokolowski, M., Broderick, J., et al. 2022, Journal of Astronomical Telescopes, Instruments, and Systems, 8, 011010, doi: 10.1117/1.JATIS.8.1.01101010.48550/arXiv.2112.00908