A CLEAN Algorithm Based on Multi-Scale Band-Pass Filtering and GPU Implementation Postprint

Authors: Mei Ying, Deng Hui, Zhang Xiaoli, Wang Feng, Chen Tengda

Date: 2018-05-15T00:00:00+00:00

Abstract

China' s new-generation radio spectral heliograph—the Mingantu Spectral Radioheliograph (MUSER)—operates with high temporal, spatial, and frequency resolution in the 0.4 GHz–15 GHz range, opening a new window for research into the physical processes of the initial energy release region in solar eruptive activities and solar electron acceleration. High-performance, high-quality solar imaging algorithms constitute a crucial research topic for the MUSER data processing pipeline. Drawing on the data processing methods of the solar interferometer array at the Paris Observatory (Meudon), this paper systematically discusses and analyzes the Multi-Scale CLEAN algorithm, presents parameters for the Multi-Scale CLEAN algorithm suitable for MUSER, and emphasizes GPU parallelization of the algorithm. Experimental results demonstrate that the improved Multi-Scale CLEAN achieves nearly a 3× improvement in algorithmic efficiency compared to the GPU-implemented Högbom CLEAN, effectively enhancing the performance of the entire MUSER data processing pipeline.

Full Text

A CLEAN Algorithm Based on Multi-Scale Bandpass Filtering and Its GPU Implementation

Ying Mei[1], , Hui Deng[3], Xiaoli Zhang[2],[1], Feng Wang[1],[3], , Tengda Chen[2] [1]Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China [2]Kunming University of Science and Technology, Kunming 650500, China [3]Center for Astrophysics, Guangzhou University, Guangzhou 510006, China University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: China' s new generation radio telescope—the Mingantu Spectral Radioheliograph (MUSER)—generates high-quality solar images with high temporal, spatial, and spectral resolution in the frequency range of 400 MHz to 15

GHz, providing a new observational window for research on solar activity and electron acceleration. High-performance imaging algorithms are essential for MUSER’ s massive data processing pipeline. Referring to the data processing methods of the Meudon Observatory in France, this paper presents a detailed analysis of multi-scale CLEAN algorithms, proposes appropriate parameters for MUSER, and describes a GPU-based implementation. Experimental results demonstrate that the improved multi-scale CLEAN achieves nearly three times the efficiency of the previously implemented Högbom CLEAN, significantly enhancing the performance of the entire MUSER data processing pipeline.

Keywords: Multi-Scale CLEAN; Radio Interferometric Imaging; GPU

Radio observations are crucial for studying solar explosive activities. The Mingantu Spectral Radioheliograph (MUSER), a spiral antenna array built in China, performs high-resolution imaging of the sun at 584 frequency points from 400 MHz to 15 GHz. Its observations fill the gap in the centimeter-decimeter wave band left by instruments such as the Nobeyama Radioheliograph (2 frequencies at 17 GHz and 34 GHz), the Nançay Radio Heliograph (5 frequencies from 150-450 MHz), and the Irkutsk Radio Heliograph (5.7 GHz), playing an important role in solar activity research and its impact on Earth [1].

MUSER consists of a low-frequency array (MUSER-I, 400 MHz–2 GHz) and a high-frequency array (MUSER-II, 2 GHz–15 GHz). Its high temporal, spatial, and frequency resolution poses significant challenges for the data processing system. The data receivers for both arrays capture observation data comprising 16 channels every 3 ms. In subsequent processing, efficiently obtaining high-quality solar images becomes the key challenge.

The main steps in radio interferometric imaging include weighting, gridding, Fourier transformation, and deconvolution. Deconvolution, the process of cleaning the dirty image, is the most time-consuming component. For solar imaging, the deconvolution process is also the most critical concern in pipeline design. Based on a thorough survey of related work, this paper further discusses multi-scale deconvolution methods in radio interferometric imaging and presents corresponding implementation approaches.

1. Related Work

The fundamental principle of radio interferometric imaging is that the actual sky brightness distribution represents the Fourier transform of the visibility function. Due to incomplete UV coverage in practical observations, the sampling function is discrete. Since Högbom proposed the CLEAN algorithm in 1974, it has seen tremendous development and widespread application in image reconstruction for radio telescopes. The key idea of the classical Högbom CLEAN algorithm is to provide a model for the actual sky brightness, iteratively searching for peak brightness positions and accumulating them as shown in equation (1), where the parameter represents the peak brightness during iteration, represents the peak position determined from the residual image’ s maximum brightness, and the iteration count is marked [2-3]. In equation (2), represents the dirty image brightness, represents the dirty beam, represents the accumulated brightness calculated from equation (1), and represents the residual image brightness.

Several variants of the Högbom CLEAN algorithm, including the Clark, Cotton-Schwab algorithms, and the Maximum Entropy Method (MEM), were subsequently proposed. The Steer Clean algorithm, used for imaging with the Nobeyama Radioheliograph, achieves cleaning efficiency nearly ten times that of the Högbom algorithm [4]. However, the basic Högbom CLEAN algorithm remains widely used due to its advantages in handling both point sources and extended sources.

With the continuous development of new-generation radio interferometers, multi-scale approaches have become one method for improving efficiency while maintaining the simplicity and signal-to-noise ratio of the basic CLEAN algorithm. Multi-scale algorithms such as Multi-Resolution CLEAN, Multi-Scale Maximum Entropy, and Wavelets CLEAN have effectively improved cleaning efficiency but suffer from fixed scale sizes and high computational costs.

Cornwell’ s Multi-Scale CLEAN algorithm, proposed in 2008, is considered direct, stable, and well-convergent [3]. To reduce iteration counts and improve algorithm stability, Multi-Scale CLEAN convolves the dirty image with multiple convolution kernels of different scales, generating a series of convolved images and searching for the global brightness peak across all images. The basic workflow is: (1) select different scale sizes, convolve with the dirty image to obtain a series of convolved images; (2) find the global maximum brightness point, recording its position and scale; (3) compute the convolution of the selected scale with the dirty beam, multiply by a gain factor, and store the result; (4) subtract the result from step (3) from all convolved images from step (1); (5) repeat steps 1-4 until reaching the loop threshold or preset iteration count; (6) convolve the stored results from step (3) with the clean beam and add the residual image to obtain the final clean image.

To control iteration counts and improve image quality, Multi-Scale CLEAN requires assigning different weights to convolved images at different scales. After each iteration, the residual image must be multiplied by its scale’ s weight to adjust amplitude. Through extensive experiments, the weight in equation (3) has proven to be a suitable parameter, where represents the weight, represents the current scale size, represents the maximum scale size, and represents the weight parameter.

To reduce the algorithm’ s dependence on gain, Cornwell selected a parabolic function as the Multi-Scale CLEAN convolution kernel, as shown in equation (4), where represents the spherical wave function, (x, y) are pixel coordinates, and represents the scale size. When in the equation, corresponds to the Dirac delta function in Högbom CLEAN. Since Gaussian convolution kernels differ only slightly from the kernel function in equation (4) at high dynamic ranges, Cornwell generally adopts Gaussian kernels of different scales. Using the clean beam size as a reference, the kernel sizes typically follow a geometric progression: 1×, 2×, 4×, and 8× the clean beam size.

To further improve processing speed, graphics processor technology has become the primary method for accelerating deconvolution in recent years. For example, GPU-based gridding algorithms have improved w-projection efficiency by nearly 100 times compared to CPU implementations [5]. Our team previously achieved more than tenfold improvement in Högbom CLEAN performance using GPU acceleration. However, parallel implementation of Cornwell’ s Multi-Scale CLEAN presents significant challenges. While the maximum search process can be parallelized, the identified maximum must be deconvolved simultaneously across all scale images, requiring cross-scale data processing that impacts parallel efficiency.

2. Multi-Scale CLEAN Based on Bandpass Filtering

2.1 Basic Principle

Referencing Cornwell’ s cleaning method, the key to accelerating the cleaning process lies in enabling simultaneous cleaning operations. In reality, due to the fundamental principle of Högbom CLEAN, the cleaning iteration process is inherently serial. The only possibility for parallelism is to independently perform multiple cleaning iterations simultaneously.

The Meudon Observatory’ s solar interferometer in France, like MUSER, is a solar-dedicated radio interferometer. Its data processing employs a multi-scale cleaning method based on frequency bandpass filtering [5, 6]. This approach relies on two simplifications: (1) no excessive extended sources exist, and (2) the spatial scale spectrum is relatively narrow. Both conditions are readily satisfied for solar observations. This method shares similarities with Cornwell’ s multi-scale cleaning but is more amenable to parallel implementation.

Drawing from the Meudon Observatory’ s approach and considering MUSER’ s antenna distribution and maximum baseline, the fundamental principles of our implemented algorithm are: (1) construct a series of continuous bandpass filters in the frequency domain uv-plane, where the filter radii grow by powers of 2 and different filters correspond to different spatial scales; unlike the French implementation, we dynamically compute the longest baseline and corresponding uv distribution for each observation, then calculate the filter passband frequencies accordingly; (2) split the original sparse UV distribution into several independent distributions at different scales; (3) for each scale’ s UV distribution, obtain the dirty image via inverse Fourier transform (FT) and perform conventional cleaning; (4) the final clean image equals the sum of clean images from all scales.

2.2 Scale Considerations

As evident from the above principles, the appropriate number of bandpass filters is critical to the algorithm. The Meudon solar interferometer has conducted preliminary studies using five filters. MUSER's data processing adopts the same approach with further improvements: based on the sun's position in altitude and azimuth, we compute the real-time UV distribution, then dynamically calculate each filter' s frequency coverage according to the farthest UV point. During MUSER imaging, five filters (F0 through F4) are used, with the UV filter design for different scales shown in [Figure 1: see original paper].

The horizontal axis represents the frequency range in units of (dynamically computed longest baseline divided by 60, where 60 is a custom parameter). Since UV coverage is symmetric in the heliograph, the UV frequency range is identical and represented by a single curve in the figure, with the corresponding UV plane shown in [Figure 2: see original paper]. Based on five different scales, the dirty image is divided into different frequency bands. As the k value increases, the scale decreases progressively until all UV values are covered. When k=0, the filter is a Gaussian filter covering 20% of the maximum UV range, representing the largest scale in imaging. For k=1, 2, 3, and 4, the filters appear as rings in the frequency domain, implementing bandpass filtering. Regardless of the number of levels, the essential requirement is that the sum of these filters must be continuous in frequency, covering all UV points, as shown in the final panel of [Figure 1: see original paper] and [Figure 2: see original paper]. Using a Gaussian filter, the synthesized filter exhibits gradually decreasing properties at the edges, which effectively reduces imaging artifacts caused by sparse UV coverage at large radii.

3. GPU Implementation

3.1 Implementation Details

MUSER' s imaging GPU implementation employs the GPU-CUDA architecture using Python with PyCUDA and Scikit-CUDA as CUDA parallel programming interfaces. Both standalone command-line and distributed operation modes have been implemented [6, 7]. Data preprocessing (format conversion [7, 8], phase calibration, anomalous data flagging, etc.) runs on the CPU, while imaging and cleaning operate on the GPU. A high-performance computing cluster comprising eight servers (CPUs+GPUs) has been deployed at the observatory. The imaging flowchart is shown in [Figure 3: see original paper], and the MUSER Multi-Scale CLEAN algorithm is described in .

Using low-frequency array observation data from "2015-11-01 04:08:49.354161240 (UTC), 1.7125 GHz, right circular polarization" as an example, the final cleaned image is shown in [FIGURE:4(a)]. CASA was used throughout experiments to verify correctness at each processing stage. Converting our system' s UVFITS files to MS format using CASA and applying multi-scale imaging yields consistent results. To demonstrate processing validity, [FIGURE:4(b)] shows published results from the Nobeyama Observatory for comparison. Since our current processing does not include P-angle correction, some errors remain in the images, which will be addressed in future work.

Table 1. Multi-Scale CLEAN for MUSER

1. Select the estimated maximum sky brightness as the iteration threshold with a gain of 0.1

a) Iterative Computation Repeat: - In parallel across all filtered images: (1) locate the brightest point, record its brightness, position, and corresponding scale, then compute its convolution with the dirty beam multiplied by the gain, storing the result; (2) subtract this result from the image to obtain the residual image; Until the maximum brightness reaches the threshold or the preset iteration count of 200 is reached

b) Obtain Clean Image For each scale: - (1) Convolve stored results with the clean beam to obtain the clean image - (2) Add the residual image to the clean image

Sum the clean images from different frequency ranges across all scales to obtain the final result.

3.2 Algorithm Efficiency

Since multiple filtered images are cleaned simultaneously, each with different iteration counts that change dynamically, we illustrate the performance of multi-scale bandpass filtering through specific examples. On an NVIDIA Corporation GM200 test platform, processing the 2015-11-01 1.7125 GHz data in [Figure 4: see original paper] using bandpass-filter-based multi-scale cleaning required 1.424 seconds, compared to 4.03748 seconds for the Högbom CLEAN algorithm from previous experiments [8, 9]. Multiple cleaning experiments demonstrate that the current Multi-Scale CLEAN effectively improves imaging efficiency by nearly three times while reducing iteration counts, representing significant progress for the MUSER data processing pipeline.

4. Conclusion

This paper analyzes the principles of Multi-Scale CLEAN algorithms and implements a bandpass-filter-based multi-scale cleaning method. Unlike Cornwell’ s multi-scale CLEAN, which iterates across global images at different scales, our approach achieves complete parallelism across scales. Experiments selected appropriate filters, scale sizes, and parameters for MUSER imaging. The GPU-based implementation demonstrates that bandpass-filter-based multi-scale cleaning achieves nearly threefold performance improvement over Högbom CLEAN under equivalent conditions while effectively reducing iteration counts. These results will significantly enhance the overall data processing system performance and provide valuable reference for radio interferometric imaging.

References

[1] Yan Yihua, Zhang Jian, Chen Zhijun, et al. Progress on Chinese spectral radioheliograph–CSRH construction[C]// Proceedings of the General Assembly and Scientific Symposium. 2011.

[2] Taylor G B, Carilli C L, Perley R A. Synthesis imaging in radio astronomy II[C]// ASP Conference Series. 1999.

[3] Cornwell T J. Multiscale CLEAN deconvolution of radio synthesis images [J]. IEEE Journal of Selected Topics in Signal Processing, 2008, 2(5): 793-801.

[4] Koshiishi H. Restoration of solar images by the Steer algorithm [J]. Astronomy & Astrophysics, 2003, 412(3): 893-896.

[5] MUSCAT D. High-Performance Image Synthesis for Radio Interferometry [J]. arXiv preprint arXiv:14034209, 2014.

[6] Mercier C, Subramanian P, Kerdraon A, et al. Combining visibilities from the giant meterwave radio telescope and the Nancay radio heliograph–High dynamic range snapshot images of the solar corona at 327 MHz [J]. Astronomy & Astrophysics, 2006, 447(3): 1189-201.

[7] Wei Shoulin, Liu Pengxiang, Wang Feng, et al. Real-time data processing in Mingantu Ultrawide Spectral Radio Heliograph based on Spark Streaming[J]. Astronomical Research & Technology, 2017, 14(4): 421-428.

[8] Chen Tairan, Wang Wei, Wang Feng, et al. The study and application of a high performance UVFITS assembly system based on MPI [J]. Astronomical Research & Technology, 2016, 13(2): 184-189.

[9] Wang F, Mei Y, Deng H, et al. Distributed data-processing pipeline for Mingantu Ultrawide Spectral Radioheliograph [J]. Publications of the Astronomical Society of the Pacific, 2015, 127(950): 383-396.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.