Exploration and Application of AI in Space Science

Xiang Zhao [*1] **You Song** [*1]

Abstract

This paper briefly introduces the application of artificial intelligence in space science. In some simulated satellite tasks in space, artificial intelligence can provide assistance when combined with traditional methods to produce better results. In the realm of satellite sampling and image reconstruction endeavors, we employ Large Language Models (LLMs) to enhance the CUDA operator within the data simulation and image reconstruction phases. This strategic application has yielded a marked improvement in operational efficiency and has effectively addressed computational bottlenecks that have historically plagued conventional methodologies. In addition, we speculate that we can use diffusion models to simulate the reconstruction of images for each period. If feasible, after the actual satellite operation time reaches the upper limit, we can approximately obtain more period reconstruction results.

1. Introduction

The exploration of the ultra-long-wavelength (ULW) radio spectrum (Boonstra et al.), specifically frequencies below 30 MHz, occupies a critical juncture in astronomical research, presenting unparalleled opportunities to probe the early universe's mysteries. This spectral range, largely uncharted due to terrestrial constraints, is where the cosmic dark ages and the epoch of reionization—key phases in cosmic history—can be studied in detail. Ground-based observations in this frequency band are significantly hindered by the Earth's ionosphere, which distorts and absorbs radio waves, and by ubiquitous radio frequency interference (RFI) from human-made sources. These challenges have necessitated a shift towards space-based observations, with the Moon's orbit emerging as a particularly promising platform. This shift not only circumvents the limitations posed by the ionosphere and RFI but also opens new avenues for acquiring high-resolution sky maps and detecting faint cosmic sources with an unprecedented level of precision.

The 1970s marked a significant era with the launch of missions like IMP-6 (Brown, 1973) and the Radio Astronomy Explorers (RAE-1 (Alexander & Novaco, 1974) and RAE-2 (Alexander et al., 1975)), which undertook the first low-frequency radio observations from space. These missions underscored the Moon's far side as an ideal location for such observations, offering a natural shield against Earth's RFI (Zheng et al.(Zheng et al., 2017)). Nevertheless, the technological limitations of that time restricted the resolution and accuracy of the collected data, highlighting the need for advancements in observational technology and strategies.

The advancement of interferometric arrays orbiting the moon signifies a paradigm shift in low-frequency radio astronomy. These arrays, using the Moon as an RFI shield and using solar power efficiently (Huang et al., 2018), promise to overcome the spatial and spectral challenges that have long hindered observations in this frequency band. The deployment of linear arrays in lunar orbit paves the way for high-precision spectral measurements and detailed celestial mappings, heralding a new era of discoveries in the least explored regions of the electromagnetic spectrum.

2. Background

The simulate image reconstruction task is used to simulate satellites that observe and obtain sensor information from the sun in lunar orbit, generate sensor data sampled by the satellite through simulation, and then generate simulation images. In this task, a mother satellite and eight sub-satellites form a satellite array that orbits the moon at an approximate inclination angle. It should be noted that the distance between satellites is not constant, and the entire satellite array orbits the moon in a spring-like stretching transformation. Lunar orbit allows sensors to sample more lossless information than the ground, but the moon itself is obscured, and not all of the sampled points can be used for subsequent missions, only about one-third of the sampling points in each period can be used for subsequent calculations.The satellite sensors for lunar orbit flight have more non-destructive information for sampling compared to the ground, but due to the obstruction of the moon itself, not all

[*]Equal contribution [1]Beihang University, School of Software, Beijing 100191, People's Republic of China. Correspondence to: You Song <songyou@buaa.edu.cn>.

Accepted at the 1st Machine Learning for Life and Material Sciences Workshop at ICML 2024.

sampling points can be used for subsequent tasks, and only about one-third of the sampling points in each period can be used for subsequent calculations.

The Imaging Algorithm for a Lunar Orbit Interferometer Array, as proposed by Huang et al. (Huang et al., 2018), is a groundbreaking approach designed to tackle the unique challenges of reconstructing sky maps from an array of satellites orbiting the Moon. This algorithm addresses the complexities introduced by the three-dimensional distribution of baselines and the time-dependent sky blockage caused by the moon's presence.

The algorithm is based on the principle that the visibility data, which are the correlations between signals from different antennas, are linearly related to the sky brightness distribution. The visibilities are measured when the satellites are on the far side of the moon, where the Earth's RFI is shielded. The data collected are then processed to reconstruct the sky map.

The algorithm begins by discretizing the sky into pixels and expressing the visibilities as a sum over these pixels (Huang et al., 2018). This is represented mathematically as:

$$V = B \cdot T + n \tag{1}$$

where V is the vector of measured visibilities, B is the response matrix that encapsulates the antenna beam patterns and the baseline distribution; T is the sky brightness temperature distribution; and n is the noise vector.

Reconstruction of the sky map T is achieved by solving the above equation, which is a linear inverse problem. The minimum variance estimator for T is given by:

$$\hat{T} = (B^H N^{-1} B)^{-1} B^H N^{-1} V \tag{2}$$

where B^H is the Hermitian transpose of B, N is the noise covariance matrix, and N^{-1} is its inverse.

The algorithm (Chen et al., 2021) also incorporates the moon's blockage into the visibility data, which is crucial for accurate sky map reconstruction. The blockage function is introduced as a shade function in the visibility equation, effectively modifying the visibility data to account for the regions of the sky that are obscured by the moon.

In this algorithm, there are mainly the following steps:

First, compute visibility.

$$V(u,v,w) = \int A(l,m)T(l,m)e^{-2\pi i[ul+vm+w(n-1)]}\frac{dldm}{n} \tag{3}$$

Then, remove the phase:

$$V'(u,v,w) = V(u,v,w) \cdot e^{-2\pi iw} \tag{4}$$

Finally, reconstruct the sky map image:

$$T_B(l,m,n) = \sum V'(u,\nu,w)e^{2\pi i(ul+vm+wn)} \tag{5}$$

In some studies, the sampling space of the baseline point is divided into some approximate subspaces, frequency statistics are performed at a low frequency, and then the rule is approximated to a higher frequency. In the process of image reconstruction, the multiplication of dots in the approximate space is added to the frequency in the approximate space to reduce the influence on the repeated sampling points in each period. The formula for this method is as follows:

$$T_B(l,m,n) = \sum V'(u,\nu,w)e^{2\pi i(ul+vm+wn)}\Delta c \tag{6}$$

where Δc is the frequency penalty of each baseline coordinate in the subspace.

However, this method has the drawback that it cannot determine whether the frequency in the low-frequency scenario still works after it is extended to the high frequency. In addition, only part of the sampling points are selected under each frequency condition, so the frequency weight on the dot multiplication of the baseline coordinates in the image reconstruction process is different from the actual one. Moreover, this frequency is the sum of all occurrences of coordinate points in the subspace, so the result is definitely inaccurate.

It is worth noting that in the task of reconstructing the sky map, we need to not only accurately reconstruct the sky map results based on satellite observation data, which requires numerical accuracy in the calculation process, but also visually observe the changes in the sky map. The former often requires a large amount of computation and is slow in speed, while the latter can provide assistance in some scenarios, such as observing the diffusion path of solar storms.

AI in Space

2.1. Optimize traditional methods

In this paper, we introduce how to use AI to complete and optimize tasks in space science. The satellite samples data on the lunar orbit, reconstructs the image on the basis of algorithms, and intuitively displays information in the data, such as the propagation of solar storms. However, we need to predict the format of satellite sampling data and the effectiveness of the algorithm in advance, so a task is needed to theoretically verify.

We break down the entire process into three parts. Firstly, there is data simulation, including baseline construction and plane table coordinate construction; next, visibility construction, which simulates the data collected by sensors; and finally, image reconstruction, which is performed to generate images based on visibility data inversion.

In the data simulation stage, we use LLMs to find a data sampling strategy that maximizes the number of points in each period and reduces repetition rates. Then, with the

assistance of GPT4, we optimized the CUDA operator by simulating sampling in batches for each satellite, as well as sorting and deduplication, so that the data simulation for one period under 10Mhz can be completed in a few minutes, as shown in Table 1. For calculating the frequency of each coordinate point appearing throughout the entire period, it is also possible to complete the statistics of billions of coordinate points in ten hours.

Table 1. Comparison of generation times using different methods in one period.

Method	Construct Baseline	Generate UVW
Matlab (1Mhz)	14.82s	58.33s
C++ Only (1Mhz)	91.82s	861.33s
Ours (1Mhz)	0.0017s	0.38s
Ours (10Mhz)	0.02s	4.53s

On this basis, we discovered the relationship between the number of simulated sampling points and the number of sampling periods at different frequencies and also found an upper bound, as shown in Figure1, which is important for future research. Researchers can predict future development and perform quantitative analysis based on this foundation.

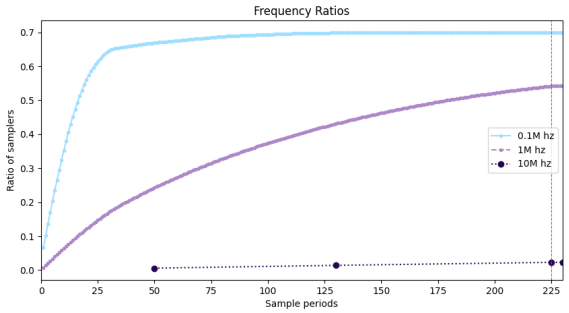


Figure 1. The relationship between the number of sampling points and the number of sampling periods. The sampling time interval is 1 second, 0.1 second, and 0.01 second from 0.1Mhz to 10Mhz respectively. In practice, satellites operate for a period of 14 days, with a maximum of 5 years, which is 130 cycles. This is the actual sampled data, but theoretically, we can reach the upper bound of the sampling points in 225 cycles (red line)

Similarly, in the simulating visibility construction stage, we can also use optimized operators, but unfortunately, the optimized operators currently cannot perform well at 10MHz, so violent operations are currently used in high-frequency situations.

In the image reconstruction stage, we use traditional algorithms to select coordinate points for reconstruction at high frequencies, as shown in the Figure 2. Specifically, since the simulated data points are in the range of -1 to 1, and only the sampling frequency changes, that is, the density of the points. To demonstrate that at 10MHz, we used uniform sampling to generate a graph (b) with the same number of points as at 1MHz. In simulation, we still use the original data.

The generated image results are shown in Figure 3. It can be seen that although similarity can be seen on the contour, there is still a lot of noise in reality, which may be related to the small proportion of sampling points in the entire space in each period. At the same time, due to the orbit of the satellite, the imbalance of sampling points is concentrated on the main orbit surface, which is the white stripe in the inversion result.

Diffusion Model in Space

From the above, it can be seen that traditional methods not only have low computational efficiency, but also require further improvement in effectiveness. The diffusion model (Ho et al., 2020; Rombach et al., 2021; Nichol & Dhariwal, 2021; Peebles & Xie, 2022) proposed in the field of generative modeling has shown promising prospects for generating high-quality images by iteratively refining noise inputs. We speculate whether diffusion models can be applied to our space task, which generate images based on previously observed periodic sequences.

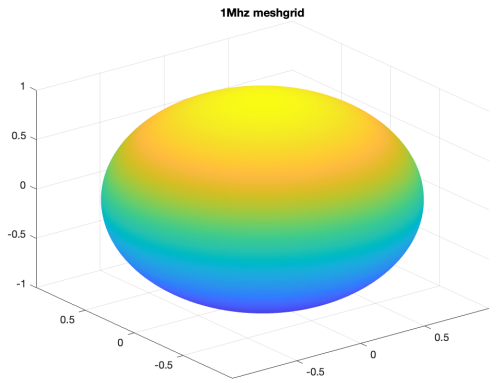
Considering image sequences as generated by diffusion processes, where the original image is the target that the diffusion model learns to reconstruct. The earlier the image in the sequence, the more noise is injected during the diffusion process, while the later the image, the less noise is injected during the diffusion process, making it closer to the original image.

We utilize an initial set of 130 periodic images, each denoted as I_t for $t = 1, 2, ..., 130$. These images represent the cumulative results up to this period t. The images are normalized to have pixel values in the range [0, 1].

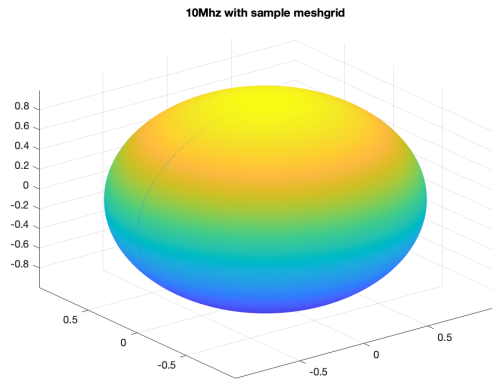
In the forward process, we define a series of noisy images x_t $t = 1, 2, ..., T$, where T is the total number of diffusion steps. Starting from the original image I_t, noise is added according to:

$$x_t = \sqrt{\bar{\alpha}_t} I_t + \sqrt{1 - \bar{\alpha}_t}\epsilon \tag{7}$$

where α_t is a predefined noise schedule, $\bar{\alpha}_t$ is the noise variance parameter of all α at step t and $\epsilon \sim \mathcal{N}(0, I)$ is Gaussian noise. The reverse process involves training a neural network ϵ_θ to predict the noise added in each step. The network is trained using a mean squared error loss between



(a) 1Mhz



(b) 10Mhz

Figure 2. Image reconstruction results. (a) is the coordinate points with a resolution of 2094 * 2094, which is 1Mhz; (b) is the coordinate points with a resolution of 10Mhz, but with uniformly distributed sampling.

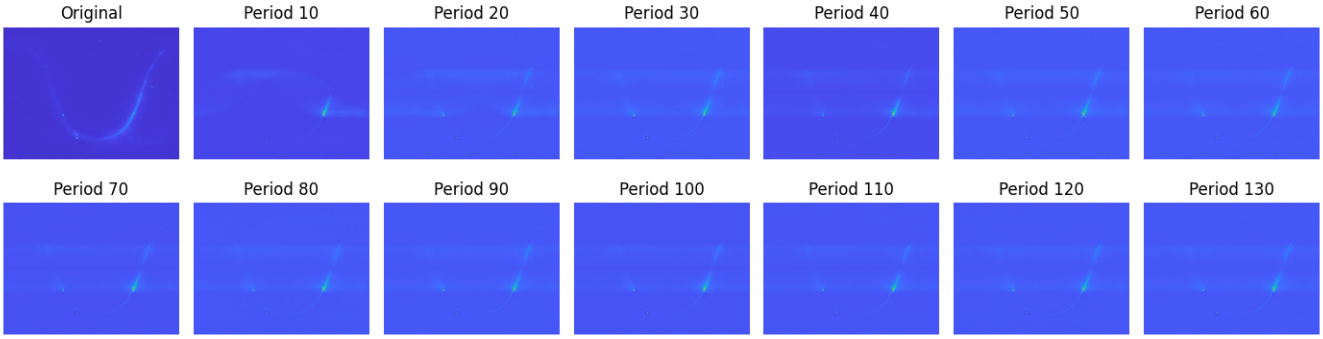


Figure 3. Results of simulating image reconstruction using traditional methods. The first image is the original image, while the others are the results of simulation generation. For example, the last image is the result of image reconstruction after simulating 130 periods.

the predicted noise and the actual noise with condition z:

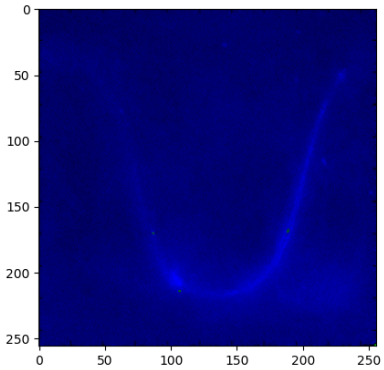
$$L(\theta) = \mathbb{E}_{t,I_t,z,\epsilon}\left[\|\epsilon - \epsilon_\theta(x_t, z, t)\|^2\right] \tag{8}$$

where x_0 is target image (idea sky image), z is normalized difference between original and noisy images, ϵ is Gaussian zero-mean unit variance noise $\mathcal{N}(0, I)$. The results obtained from the model, after filtering and trend removal, are compared as shown in Fig. 4

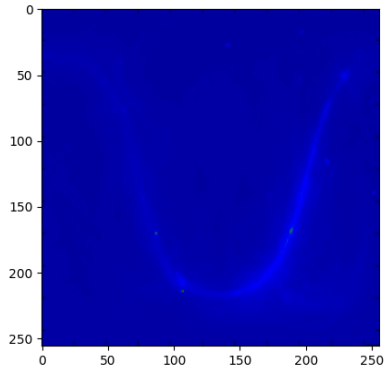
The diffusion model has enormous potential in this field, even when there are few image sequences, the model can basically show comparable contour effects. Our results demonstrate that the diffusion model effectively captures the temporal evolution of the cumulative images. The predicted images for future periods maintain the visual characteristics and cumulative patterns observed in the training set. In practical applications, the number of images generated by satellite sampling may far exceed the simulation results.

3. Discussion

This paper summarizes some simple practices using AI in space science. With the assistance of large language models, data simulation and processing can be greatly improved, which help finding latent patterns, such as the relationship between the number of sampling points and the number of sampling periods. In terms of image reconstruction, the images generated in each period need to be accumulated together to form the final result. Similarly to the image generation process of diffusion models, the time constraints faced by satellites in actual operation can also be simulated through diffusion models to approximate more steps, we have successfully applied diffusion models to predict future cumulative images in a sequential process. This approach



(a) result1



(b) result2

Figure 4. Results predicted by diffusion models using images with different levels of noise. Preprocessed image size is 256 * 256, which is similar in outline to the original image.

leverages the strengths of diffusion models in handling noisy data and iterative refinement, making it a promising tool for time-series image prediction tasks.providing a feasible solution for subsequent work. In future, we will focus on refining the model architecture and exploring its application to other domains requiring sequential image prediction. In addition, we will also investigate whether LLMs can be used to simulate satellite orbit around the moon and other scenarios, and researchers can explore more applications of AI in space science based on this research.

References

Alexander, J. K. and Novaco, J. C. Survey of the galactic background radiation at 3.93 and 6.55 mhz. *Astronomical Journal, Vol. 79, p. 777 (1974)*, 79:777, 1974.

Alexander, J. K., Kaiser, M. L., Novaco, J., Grena, F., and Weber, R. Scientific instrumentation of the radio-astronomy-explorer-2 satellite. *Astronomy and Astrophysics, vol. 40, no. 4, May 1975, p. 365-371.*, 40:365–371, 1975.

Boonstra, A., Saks, N., Falcke, H., Klein-Wolt, M., Bentum, M., Rajan, R., Wijnholds, S., Arts, M., van't Klooster, K., and Beliën, F. A low-frequency distributed aperture array for radio astronomy in space.

Brown, L. W. The galactic radio spectrum between 130 and 26ookhz. *Astrophysical Journal, Vol. 180, pp. 359-370 (1973)*, 180:359–370, 1973.

Chen, X., Yan, J., Deng, L., Wu, F., Wu, L., Xu, Y., and Zhou, L. Discovering the sky at the longest wavelengths with a lunar orbit array. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2188):20190566, 2021. doi: 10.1098/rsta.2019.0566.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *ArXiv*, abs/2006.11239, 2020. URL https://api.semanticscholar.org/CorpusID:219955663.

Huang, Q., Sun, S., Zuo, S., Wu, F., Xu, Y., Yue, B., Ansari, R., and Chen, X. An Imaging Algorithm for a Lunar Orbit Interferometer Array. , 156(2):43, August 2018. doi: 10.3847/1538-3881/aac6c6.

Nichol, A. and Dhariwal, P. Improved denoising diffusion probabilistic models. *ArXiv*, abs/2102.09672, 2021. URL https://api.semanticscholar.org/CorpusID:231979499.

Peebles, W. S. and Xie, S. Scalable diffusion models with transformers. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4172–4182, 2022. URL https://api.semanticscholar.org/CorpusID:254854389.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10674–10685, 2021. URL https://api.semanticscholar.org/CorpusID:245335280.

Zheng, H., Tegmark, M., Dillon, J. S., Kim, D. A., Liu, A., Neben, A. R., Jonas, J., Reich, P., and Reich, W. An improved model of diffuse galactic radio emission from

10 MHz to 5 THz. , 464(3):3486–3497, January 2017.
doi: 10.1093/mnras/stw2525.