PolarRec: Improving Radio Interferometric Data Reconstruction Using Polar Coordinates

Supplementary Material

This document provides supplementary information that is not included in our main paper due to space limitation: Section A further explains the visibility distance in our method. Section B describes the details of our experiments. Section C presents some supplementary experimental results.

A. Additional Explanation of Visibility Distance

The visibility distance is between real and predicted visibility in complex values, reflecting the amplitude and phase information of visibility data.

In radio interferometry imaging process, amplitude represents the intensity or energy of a signal, and corresponds to the strength with which an image responds to the 2D sinusoidal wave at a specific frequency [5]. Phase encodes the positional information of signal components [6], helping reconstruct the spatial relationships and structures within the image data.

The magnitude of a vector in a 2D plane, given by $\sqrt{A^2+B^2}$, corresponds to the absolute value of the complex number. The magnitude of $\vec{p_r}$ is $|\vec{p_r}| = \sqrt{A_r^2+B_r^2}$, which is the same as the amplitude of $V_r(u,v)$. Similarly, The magnitude of $\vec{p_p}$, $|\vec{p_p}| = \sqrt{A_p^2+B_p^2}$ represents the amplitude of $V_p(u,v)$. The phase of a complex number $A+iB$ is the angle between the vector and the real axis in the complex plane, typically measured in radians or degrees. This angle can be calculated using the arctangent function: $\theta = \arctan\left(\frac{B}{A}\right)$. For $\vec{p_r}$, the angle $\theta_r = \arctan\left(\frac{B_r}{A_r}\right)$ represents the phase of $V_r(u,v)$. For $\vec{p_p}$, $\theta_p = \arctan\left(\frac{B_p}{A_p}\right)$ corresponds to the phase of $V_p(u,v)$.

The magnitude of the vectors $|\vec{p_r}|$ and $|\vec{p_p}|$ represent the amplitudes of the complex visibility values, reflecting the strength or intensity of the signal components in the spatial frequency domain. The angles θ_r and θ_p of the vectors correspond to the phases of the complex visibility values, indicating the positional information and structural arrangement of the image components. As a result, the single point visibility distance defined by $d(\vec{p_r}, \vec{p_p})$ covers information from both amplitude and phase.

B. Additional Experiment Details

B.1. Implementation Details of PolarRec

We provide our demo code in the supplementary material. All the code we used for the experiments will be public if the paper is published. In our implementation of PolarRec, we use a 2-layer MLP with a Leaky ReLU activation in the intra-group encoder, followed by an adaptive average pooling layer. The inter-group encoding is done by a Transformer encoder. We then use an 8-layer MLP in the conditioned neural field, and only the first 8 output tokens with the dimension of 1024 from the Transformer encoder are used to condition this 8-layer MLP. We show the details of the FiLM conditioning network and PorlarRec encoder in our model in table 1 and table 2.

Our model is trained with Adam with $lr = 0.0001, \beta_1 = 0.9, \beta_2 = -0.999$, eps $= 1 \times 10^{-8}$, weight decay $= 0$. The two scaling factors α and β in the Radial Visibility Loss are both set to 1. The model for overall comparison is trained with batch_size = 32 and group_size = 32.

B.2. Baseline Details

B.2.1 CLEAN

CLEAN [4] facilitates the extraction of the original image, denoted as $I(l,m)$, from the observed visibilities $V(u,v)$ and the telescope array configuration $W(u,v)$. This process can be mathematically expressed as a Fourier inversion of the product of $V(u,v)$ and $W(u,v)$, leading to the equation $F^{-1}[V(u,v)W(u,v)] = I(l,m) * F^{-1}[W(u,v)]$. The specific arrangement of the telescope array, $W(u,v)$, yields a point-spread function (PSF) in the image plane, known as the *dirty beam*, represented by $F^{-1}[W(u,v)]$. The initial image created directly from the complex visibilities, often referred to as the *dirty image*, is the convolution of the dirty beam with the true celestial image $I(l,m)$ [9]. The process of deconvolution in the CLEAN algorithm involves an iterative technique where the peak emission in the dirty image, once convolved with the dirty beam, is systematically subtracted. Throughout this iterative process, a model of the *clean* emission is progressively constructed. In the implementation, we set the threshold to stop iteration to 1×10^{-11}, gain $= 0.1$, beam_size $= 4$, maxIteration $= 1 \times 10^5$ for the best performance.

Component	Details
Positional Embedding	PE_Module: Fourier Encoding
Layers (FiLMLinear x7)	Linear: in_features=258 (for the first layer), 256 (for subsequent layers), out_features=256, bias=True (first layer only) Activation1: LeakyReLU (negative_slope=0.01) Activation2: LeakyReLU (negative_slope=0.01) Film1: Linear, in_features=1024, out_features=256, bias=True Film2: Linear, in_features=1024, out_features=256, bias=True
Activations	ModuleList containing 7 ReLU activations
Final Layer (FiLMLinear)	Linear: in_features=256, out_features=2, bias=True Activation1: LeakyReLU (negative_slope=0.01) Activation2: LeakyReLU (negative_slope=0.01) Film1: Linear, in_features=1024, out_features=2, bias=True Film2: Linear, in_features=1024, out_features=2, bias=True

Table 1. FiLM Conditioning Component Structure

Component	Details
Linear Embedding for Visibility Values	Sequential (Linear: in_features=2, out_features=254, bias=True)
Intra-group Encoding	Sequential (Linear: in_features=512, out_features=256, bias=True, LeakyReLU: negative_slope=0.01, Linear: in_features=256, out_features=512, bias=True) AdaptiveAvgPool2d, output_size=(1660 // group_size, 512) Dropout, p=0.0
Inter-group Encoding	Transformer with 4 sets of module lists, each containing: - Residual with PreNorm and Attention (to_qkv Linear: in_features=512, out_features=1536, bias=False, to_out Sequential: Linear in_features=512, out_features=512, bias=True) - Residual with PreNorm and FeedForward (Sequential: Linear in_features=512, out_features=512, bias=True)
Output Token Heads	Module list containing 8 Sequential modules, each with: - LayerNorm: (512,), eps=1e-05, elementwise_affine=True - Linear: in_features=512, out_features=1024, bias=True

Table 2. PolarRec Encoder Component Structure

B.2.2 U-Net

In the implementation of U-Net, we follow the source code of U-Net with attention and residual blocks for MRI reconstruction [10]. The image size is set to 256 and the number of residual blocks is 2. The attention resolution is set to 20 and the number of heads is set as 4 for the best performance.

B.2.3 Radionets

We follow the original source code for Radionets [7] implementation. We choose arch_name = 'SRResNet' as the backbone network for the best performance and set amp_phase $= false$ to output the real and imaginary value of visibility data for comparison.

B.2.4 Neural Interferometry

We use the original code of Neural Interferometry [9]. The batch size is set to 4 due to the GPU memory capacity limit and the loss type is set as 'spectral' for the best performance. All other settings are the same as the default values in the original code.

Table 3. Dataset Information.

Dataset	Total Size	Test Set Size
MG	1853	370
IRSG	2027	405
UTSG	1829	365
EGB	1873	374

Table 4. Overall performance comparison on Galaxy10 Dataset [3].

Models	Domain		Metrics		
	Img	Vis	LFD↓	PSNR↑	SSIM↑
Dirty			N/A	10.204 ± 1.079	0.6583 ± 0.0543
CLEAN [4]	✓		N/A	17.535 ± 2.195	0.8023 ± 0.0304
U-Net [10]		✓	1.465 ± 0.318	16.175 ± 1.952	0.7814 ± 0.0323
Radionets [7]		✓	1.173 ± 0.262	19.575 ± 2.350	0.8305 ± 0.0300
Neural Interferometry [9]		✓	0.962 ± 0.320	22.956 ± 2.617	0.8785 ± 0.0302
PolarRec		✓	$\mathbf{0.658 \pm 0.243}$	$\mathbf{26.225 \pm 2.751}$	$\mathbf{0.9002 \pm 0.0268}$

B.3. Dataset Details

We evaluated the methods on 4 datasets, Merging Galaxies (MG), In-between Round Smooth Galaxies (IRSG), Unbarred Tight Spiral Galaxies (UTSG), and Edge-on Galaxies with Bulge (EGB). For each dataset, we randomly split 20% of all images for testing, with the remainder being used for training. The details of the datasets are shown in Table 3. Each observation of an image has 1660 sampled visibility points and all the images in our experiments are converted to grayscales for sky intensity information and then scaled to a size of 256×256. Following the methods of Wu et al. [9], we apply the inversed discrete Fourier transform (IDFT) technique to create dirty images out of the visibility data.

C. Additional Experiments

We also evaluate the models on Galaxy10 dataset [3] which contains 10 kinds of distinct galaxy morphologies. This dataset comprises 17,736 galaxy images, sourced from the DESI Legacy Imaging Surveys [2]. This in turn, merges data from the Beijing-Arizona Sky Survey (BASS) [11], the DECam Legacy Survey (DECaLS) [1], and the Mayall z-band Legacy Survey [8]. The data synthesis process on this dataset is the same as other four datasets in our paper. We then randomly split 5000 images for testing, with the remainder being used for training. The overall performance comparison results are shown in Table 4.

References

[1] Robert D Blum, Kaylan Burleigh, Arjun Dey, David J Schlegel, Aaron M Meisner, Michael Levi, Adam D Myers, Dustin Lang, John Moustakas, Anna Patej, et al. The decam legacy survey. In *American Astronomical Society Meeting Abstracts# 228*, pages 317–01, 2016. 3

[2] Arjun Dey, David J Schlegel, Dustin Lang, Robert Blum, Kaylan Burleigh, Xiaohui Fan, Joseph R Findlay, Doug Finkbeiner, David Herrera, Stéphanie Juneau, et al. Overview of the desi legacy imaging surveys. *The Astronomical Journal*, 157(5):168, 2019. 3

[3] Leung Henry. Galaxy10 decals dataset. https://github.com/henrysky/Galaxy10, 2021. 3

[4] JA Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 15:417, 1974. 1, 3

[5] Liming Jiang, Bo Dai, Wayne Wu, and Chen Change Loy. Focal frequency loss for image reconstruction and synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 13919–13929, 2021. 1

[6] Alan V Oppenheim and Jae S Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, 1981. 1

[7] Kevin Schmidt, Felix Geyer, Stefan Fröse, Paul-Simon Blomenkamp, Marcus Brüggen, Francesco de Gasperin, Dominik Elsässer, and Wolfgang Rhode. Deep learning-based imaging in radio interferometry. *arXiv preprint arXiv:2203.11757*, 2022. 2, 3

[8] David R Silva, Robert D Blum, Lori Allen, Arjun Dey, David J Schlegel, Dustin Lang, John Moustakas, Aaron M Meisner, Francisco Valdes, Anna Patej, et al. The mayall z-band legacy survey. In *American Astronomical Society Meeting Abstracts# 228*, pages 317–02, 2016. 3

[9] Benjamin Wu, Chao Liu, Benjamin Eckart, and Jan Kautz. Neural interferometry: Image reconstruction from astronomical interferometers using transformer-conditioned neural fields. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022. 1, 2, 3

[10] Yutong Xie and Quanzheng Li. Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VI*, pages 655–664. Springer, 2022. 2, 3

[11] Hu Zou, Xu Zhou, Xiaohui Fan, Tianmeng Zhang, Zhimin Zhou, Jundan Nie, Xiyan Peng, Ian McGreer, Linhua Jiang, Arjun Dey, et al. Project overview of the beijing–arizona sky survey. *Publications of the Astronomical Society of the Pacific*, 129(976):064101, 2017. 3