

基于深度学习的SKA图像反卷积研究

卢梅¹, 张利^{1*}, 李丹宁², 米立功³, 刘祥⁴, 张明⁴, 贺春林⁵,
潘伟⁵, 王蓓¹

¹贵州大学大数据与信息工程学院, 贵州 贵阳

²贵州省科学院, 贵州 贵阳

³黔南民族师范学院物理与电子科学学院, 贵州 都匀

⁴中国科学院新疆天文台, 新疆 乌鲁木齐

⁵西华师范大学计算机学院, 四川 南充

收稿日期: 2021年12月26日; 录用日期: 2022年1月23日; 发布日期: 2022年1月30日

摘要

干涉测量使得观测天文图像的分辨率显著提升, 然而其阵型所带来的点扩展函数效应需要图像反卷积技术来消除。尽管传统CLEAN反卷积算法已经广泛应用于射电天文图像的点扩展函数消除, 但仍然存在精度不高的问题。为了解决国际大科学工程——平方公里阵(SKA)的图像模糊问题, 本文提出一种深度卷积神经网络来提升射电天文图像重建的质量。实验显示, 相较于通用的方法, 本文提出的方法能更好地重建弱源, 并在整体图像质量上有明显提升。

关键词

SKA, 深度学习, 图像反卷积, 射电天文

Study on SKA Image Deconvolution Using Deep Learning

Mei Lu¹, Li Zhang^{1*}, Danning Li², Ligong Mi³, Xiang Liu⁴, Ming Zhang⁴, Chunlin He⁵,
Wei Pan⁵, Bei Wang¹

¹College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

²Guizhou Academy of Sciences, Guiyang Guizhou

³School of Physics and Electronics, Qiannan Normal University for Nationalities, Duyun Guizhou

⁴Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi Xinjiang

⁵Computer School of China West Normal University, Nanchong Sichuan

Received: Dec. 26th, 2021; accepted: Jan. 23rd, 2022; published: Jan. 30th, 2022

*通讯作者 Email: lizhang.science@gmail.com

文章引用: 卢梅, 张利, 李丹宁, 米立功, 刘祥, 张明, 贺春林, 潘伟, 王蓓. 基于深度学习的 SKA 图像反卷积研究[J]. 应用数学进展, 2022, 11(2): 613-620. DOI: 10.12677/aam.2022.112068

Abstract

Radio interferometry makes the observing resolution of astronomical images significantly improved, but deconvolution is required to eliminate the effects of the point spread function (PSF). Although the traditional CLEAN-based deconvolution has been widely used to eliminate the observed PSF, it still has the problem of low accurate reconstruction. To solve the PSF problem of the square kilometer array (SKA), a deep convolutional neural network is proposed to improve the quality of radio image reconstruction. Experiments show that compared with the traditional method—CLEAN, the method proposed in this paper can better reconstruct weak sources and significantly improve the quality of an image.

Keywords

SKA, Deep Learning, Image Deconvolution, Radio Astronomy

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

干涉测量通过联合多个小口径望远镜/天线来获得对天体的高分辨率观测[1] [2], 极大地促进射电天文学的发展[3]。诸如 Jansky 甚大干涉阵(JVLA)和阿塔卡马大型毫米波/亚毫米波阵列(ALMA)等大量的干涉阵已经在天文界得到成熟的发展。平方公里阵(SKA)是正在建造的中国参与的国际大科学工程装置, 其接受面积达到公里级, 具有超高的灵敏度、分辨率等优点[4]。SKA 极其优良的观测能力有望在宇宙起源等基本的天体物理问题上带来变革的同时, 也带来了巨大的数据处理挑战[5] [6]。

干涉测量有效地解决了大口径望远镜因技术和成本等因素难以建造的难题, 也显著提升了观测的分辨率, 使得更多的天体物理现象能够被揭示。然而由于天线数量和分布等因素的影响, 使得空间频率无法获得满足奈奎斯特定理的采样。不完全的采样会使得测量的天体图像变得模糊,

$$I^{dirty} = I^{sky} * B \quad (1)$$

其中, I^{dirty} 表示观测的模糊图像, I^{sky} 表示干净的天体图像, B 表示点扩展函数, $*$ 表示卷积计算。从模糊的观测图像中恢复出干净的天体图像一直是一个具有挑战性的问题[7]。不完全和非规则的采样形成的带有主瓣和大量旁瓣的点扩展函数(也称为脏束, dirty beam)是观测图像变得模糊的主要因素之一。如何能够有效地消除点扩展函数的模糊影响也是其关键技术之一。

消除点扩展函数模糊影响的方法通常称为反卷积。该领域经过几十年的发展, 已经发展出大量的方法, 如最大熵方法(Maximum Entropy Methods, MEM)、压缩感知(Compressive Sensing, CS)和洁化(CLEAN) [8]。MEM 是一类包含显式优化的方法, 并且结果受到熵函数的约束。熵函数是一个控制模型图像的函数, 通常具有平滑先验。因此 MEM 算法重建的图像通常更为平滑。CS 方法与 MEM 有些相似, 都使用正则化方法来优化结果, 区别在于它们的理论基础和使用的正则化不同。CS 方法基于稀疏表示理论和随机测量。CS 理论显示如果信号能够被稀疏表示在满足一定测量条件下, 只需要少量的采样即可完全恢复出信号。这突破了奈奎斯特采样定理, 提供一个新的测量和重建范式。然而, 天文测量常常无法满足 CS

的测量条件,但这也限制它在天文领域里的应用。CLEAN 方法是一类基于函数分解的贪婪算法。由于其原理简单、易于实现和易于与其它方法结合等特点,该方法在射电天文领域的应用最为广泛。

原始的 CLEAN 算法由瑞典科学家 Högbom 在 1974 年提出[9],已经对射电天文发展产生了巨大的推动作用[10]。在干涉测量发展初期,由于望远镜的灵敏度有限,测量的天文源大多是致密的,因此,该算法主要是针对致密源设计的。它将射电天空表示成点源 δ 函数的集合。该方法首先找到观测图像中的最可信的分量位置,即最大点,然后在该位置消除其点扩展函数的影响。为减少反卷积过程的震荡并加速收敛,通常会使用循环增益来修正幅度。CLEAN 算法在处理位置上很好分开的点源也很有效,但在处理延展源时性能欠佳。

处理延展源时,CLEAN 反卷积残差和模型中常常出现条纹。这是由不精确的物理表示造成的。因此,大量的 CLEAN 变种被提出来用于解决这个问题,如多尺度 CLEAN [11]。多尺度 CLEAN 使用能够表示多个尺度的基函数,使得 CLEAN 能够表示多尺度的射电天体结构。多尺度 CLEAN 使用多个用户指定的尺度来表示射电天体结构,有效地解决了基于点源分解所带来的反卷积条纹。大量的演化算法也被提出来以解决不同场景的任务[8] [12] [13]。

CLEAN 算法及其变种属于模型驱动的方法,这类方法的主要缺点是算法的性能常常随着场景和参数改变而发生性能的显著变化。数据驱动方法通过从海量数据中学习规律,能够有效地解决模型驱动方法所带来的缺陷。本文提出一种数据驱动的方法用于 SKA 点扩展函数的消除来恢复高动态范围的图像。

第二节将仔细地介绍本文的算法,包括损失函数、随机梯度下降、归一化和网络架构等。第三节将使用 SKA 模拟数据来验证本文算法的性能,并与多尺度 CLEAN 算法进行比较分析。第四节将总结本文工作。

2. 本文算法

在这节中,我们描述本文提出的深度卷积神经网络。受 Zhang 2017 [14]启发并进一步探索批归一化的思想,本文提出一个新的深度卷积神经网络(本文称之为 DCN),同时也用了批归一化来解决内部协变量漂移问题,加速网络收敛。

2.1. 表示

我们需要解决的问题是消除 SKA 点扩展函数所带来的模糊影响。在 CLEAN 算法及其变种中,射电天体结构是通过在观测图像中不断搜寻和分离模型分量来实现的,即模型图像 I^{model} 表示基函数的集合:

$$I^{model} = \sum_1^n I_i^{comp} \quad (2)$$

其中, I_i^{comp} 表示模型分量, n 是用于表示模型图像的分量数目。对于原始 CLEAN 算法, I_i^{comp} 是 δ 函数 $\delta\{x, y\}$ ($\{x, y\}$ 表示二维位置信息)。对于多尺度 CLEAN 算法和自适应尺度 CLEAN 算法, $I_i^{comp} = f(x, y, z)$, 其中, z 表示幅度, f 为基函数,通常是高斯函数等[7]。本文方法则与 CLEAN 及其变种的思想不同,通过海量数据学习方式来获得 B 的逆(B^{-1})。 B^{-1} 是通过神经网络节点的方式组织在一起的。干净天体图像可以通过脏图与 B^{-1} 卷积获得,

$$I^{model} = I^{dirty} * B^{-1} + \varepsilon \quad (3)$$

其中, ε 是天空模型 I^{model} 和干净天空图像 I^{sky} 的误差。当网络收敛时,这应该是一个无穷小量。

2.2. 损失函数

损失函数用于训练过程中梯度的计算,从而有效地指导训练的方向。在 DCN 训练中,我们使用均方根误差来判断重建的图像与输出参考图像之间的差异性,

$$L(\theta) = \frac{1}{2N} \sum_{i=1}^N \left\| (I^{dirty}, \theta) - I^{clean} \right\|^2 \quad (4)$$

其中 $L(\theta)$ 为损失函数, θ 为 DCN 学习的参数, N 为 dirty-clean 图像对的数目。

2.3. 优化器

Adam 优化算法广泛应用在深度学习的训练和优化, 在计算机视觉等领域的应用获得广泛成功。它是一种基于随机梯度下降的优化方法。Adam 算法策略可以表示为:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (5)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (6)$$

$$M_t = \frac{m_t}{1 - \beta_1^t}, \quad V_t = \frac{v_t}{1 - \beta_2^t} \quad (7)$$

$$W_{t+1} = W_t - \frac{\varphi}{\sqrt{v_t + \varepsilon}} M_t \quad (8)$$

其中, g_t 表示随机目标函数 $f_t(W)$ 的梯度, m_t 表示一阶矩, v_t 表示二阶矩, $\beta_1, \beta_2 \in [0, 1)$ 表示矩估计的指数衰减率。 φ 为学习率, 表示参数空间有效步长的量级, $\omega = 10^{-8}$ 代表一个常数参量。衰减过程分别对一阶矩 m_t 、二阶矩 v_t 偏差校正, 得到 M_t 、 V_t 。每经过一次衰减的迭代后都需要更新随机目标函数 $f_t(W)$ 的 W 参数。

经典的随机梯度下降方法的学习率在整个训练过程中是保持不变的, 即所有的权重更新所使用的学习率是相同的。Adam 与随机梯度下降法不同的是, 该方法的学习率不再是固定的, 而是自适应的, 它通过梯度的第一次和第二次矩来计算和更新学习速率。

2.4. 批归一化

批归一化是一个神经网络训练的技巧, 不仅可以加快模型的收敛速度, 还可以在在一定程度上缓解网络中的“梯度弥散”问题, 从而使训练的深层网络模型更加稳定。在训练网络过程中, 参数不断发生更新, 除了输入层的数据外, 后面网络每一层的输入数据分布随着参数更新, 也不断发生变化。具体过程如下:

选取小批量数据, 然后计算小批量数据的均值和方差,

$$\mu_\beta \leftarrow \frac{1}{m} \sum_{i=1}^m x^{(i)} \quad (9)$$

$$\sigma_\beta^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x^{(i)} - \mu_\beta)^2 \quad (10)$$

其中, μ_β 表示每一个训练批次数据的均值, σ_β^2 表示每一个训练批次数据的方差, m 为 batch 大小, i 为训练批次。

接着进行归一化处理, 获得 0-1 分布,

$$X_i \leftarrow \frac{x^{(i)} - \mu_\beta}{\sqrt{\sigma_\beta^2 + \varepsilon}} \quad (11)$$

$$y_i \leftarrow \gamma X_i + \beta = BN_{\gamma, \beta}(x_i) \quad (12)$$

其中, γ 是尺度因子, β 是平移因子, 是在训练时网络自己学习得到的。参数依赖于梯度的更新, 数据分布不同往往出现梯度爆炸或消失的问题。内部协变量漂移[15]是造成该问题的主要原因之一。在训练时, 不同数据的分布差异使得网络权重更新有较大不同, 这种不同会在层级之间传播, 使梯度出现振荡、爆炸或消

失的问题。同时,所带来的权重振荡既降低了网络训练速度,也降低了网络性能。在验证和测试时,与训练集分布不同会带来学习到的模型的泛化能力大大降低。尽管小批量的随机梯度下降已经广泛应用在 CNN 的训练,然而内部协变量漂移,大大降低了训练的有效性[15]。为避免该问题的出现,批量归一化将特征强制变换到均值为 0,方差为 1 的标准化分布。变换后的分布落在激活函数对输入比较敏感的区域,有效地防止了梯度爆炸或消失的问题。批量归一化具有诸如对初始化不敏感、更快的收敛速度和更好的重建性能等优点。

2.5. 网络架构

DCN 的输入为观测图像,即包含点扩展函数模糊效应的图像,而输出则为不包含观测效应的图像。网络学习输入与输出图像之间的映射函数。网络共有 10 层,前 8 层的每层包含卷积和激活函数两个子层;倒数第二由卷积、批归一化和激活函数三个子层组成;最后一层仅包含卷积层。10 层的网络结构能够有效地表征复杂的特征。为保证特征抽取的稳定性,正如其它研究一样,我们采用的卷积核大小为 3×3 。每一层采用 64 个节点。网络架构如图 1 所示。

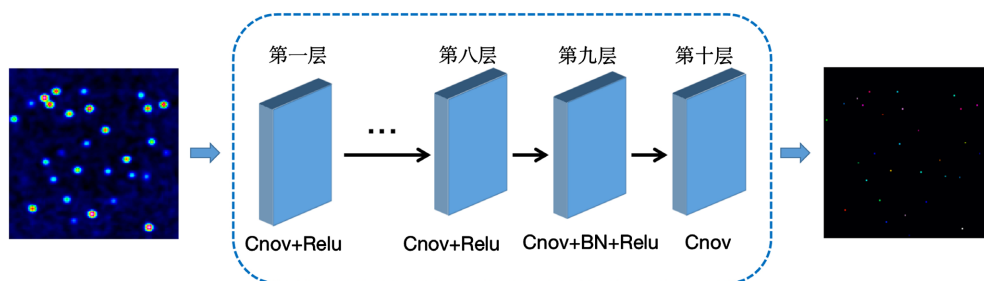


Figure 1. The network architecture of DCN
图 1. DCN 的网络架构

3. 实验与分析

SKA 具有平方公里级的合成孔径,具有超高的分辨率。SKA 第一期低频天线阵列(SKA1-low)已经开始建设,采用多站干涉。由于天线位置分布较稀疏,导致大量的空间频率域无法获得数据。正如前面分析,这会模糊天体图像,使得部分特征无法分辨,直接影响到天体物理研究。SKA1-low 的首要科学目标是宇宙再电离信号的探测,强的前景信号会淹没微弱的宇宙再电离信号,使得宇宙再电离信号重建变得非常具有挑战性。在强的前景信号中,河外点源是在巡天观测中最为常见的一种强源。河外点源的重建和消除在宇宙再电离观测数据中占有非常重要的地位。为此,本文主要的研究对象是河外点源。

为了展示 DCN 在消除点扩展函数的性能,我们利用 RASCIL 模拟了 SKA1-low 核心阵列进行模拟观测。图 2 中展示了模拟观测的一个例子。从图像可以看出,相对于干净的天体图像,观测后的脏图中许多细节结构变得模糊不清,无法辨识,进而无法进行相关的天体物理研究。

利用 DCN 算法处理的结果显示在图 3。图 3(a)为重建的模型图像,用于逼近真实的天空。图 3(b)为模型误差图像,它是重建的模型图像与参考图像的差,用于衡量重建的模型图像与真实参考图像之间差异。可以看出,重建的模型图像与参考图像在结构上非常相似。其模型误差图像的微弱特征也可以证明上述结论。整体上,DCN 能够很好的重建点源。

为进一步测试 DCN 的性能,通过模拟产生了不同分布的随机点源,并且在亮度上也具有随机性。这能够更好地模拟天空河外点源的情形。其模拟观测结果显示在图 4。在这里,我们将 DCN 与通用的 CLEAN 反卷积的结果进行了比较,如图 5,可以看出 DCN 重建的结果明显优于 CLEAN 处理的结构。整体上,DCN 不但能够很好的重建不同位置的源,而且重建的图像优于 CLEAN。

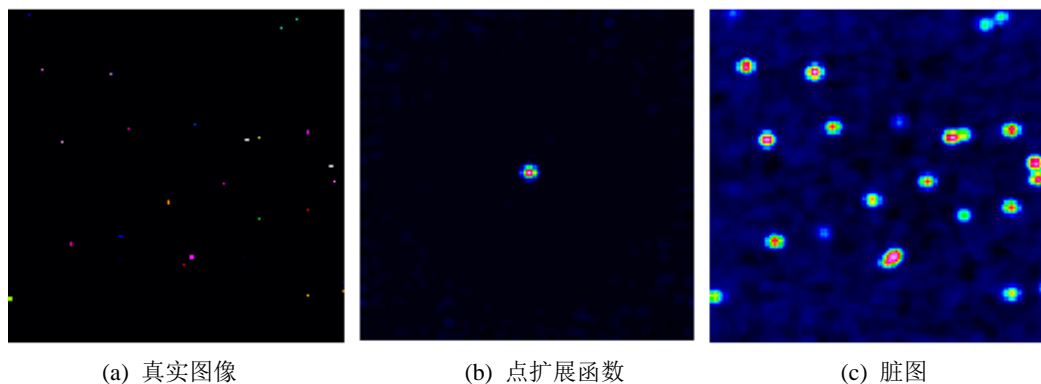


Figure 2. Simulation observations from RASCIL
图 2. RASCIL 模拟观测结果

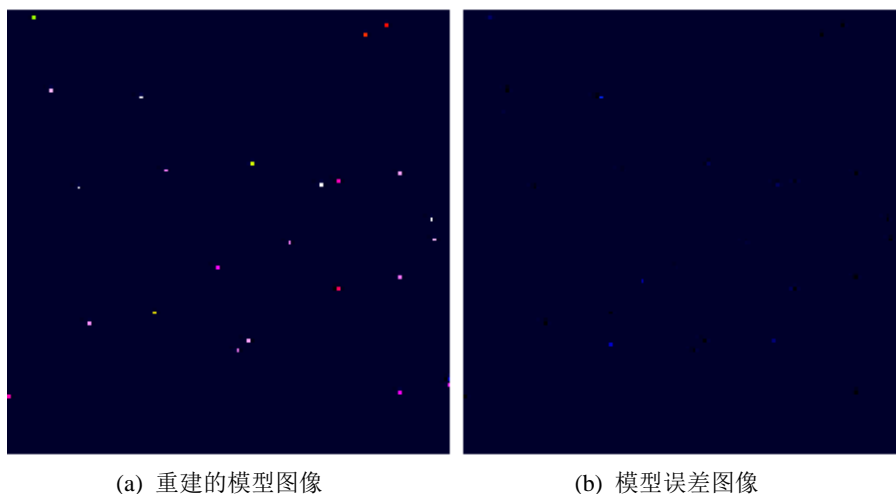


Figure 3. Results from DCN algorithm
图 3. DCN 算法处理的结果

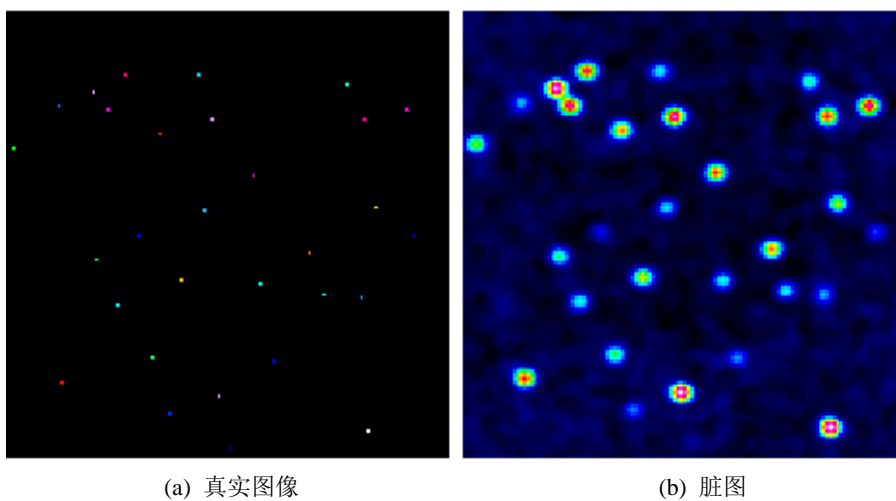


Figure 4. Simulation observations from RASCIL
图 4. RASCIL 模拟观测的结果

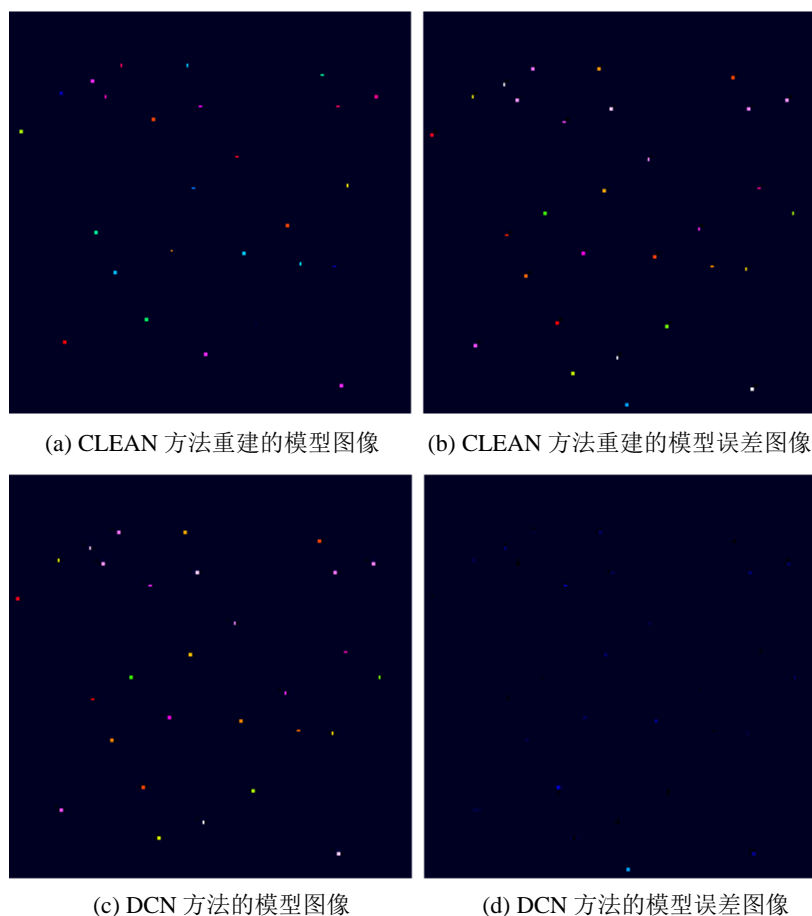


Figure 5. Comparison of results from typical CLEAN algorithm and the improved algorithm

图 5. 通用的 CLEAN 方法与本文的 DCN 方法图像重建结果对比

4. 总结

正在建设的 SKA 将是全世界最先进的射电望远镜。该望远镜使用了干涉测量技术，通过大量的小口径望远镜或天线来实现虚拟的大口径望远镜的效果。由于望远镜分布较为稀疏，部分的空间频率域无法采样，这在空间域会形成一个具有模糊效应的点扩展函数。CLEAN 算法是通用的消除点扩展函数的反卷积方法，采用模型驱动的设计理念。模型驱动方法通常会有较明显的误差。本文提出一种新的深度学习模型 DCN 用于 SKA 图像重建。该方法采用数据驱动的设计理念，使用海量数据来更好地学习数据中包含的模式。从结果上看，新方法 DCN 能够更加有效地重建 SKA 图像。

基金项目

国家自然科学基金(11963003)，SKA 专项资助(2020SKA0110300)，国家重点研发计划(2018YFA0404602)，黔科合重大专项字[2019]3011，贵州大学引进人才科研基金(贵大人基合字(2018)60号)，贵州大学培育项目(贵大培育[2020]76号)。

参考文献

- [1] Pawsey, J.L., Payne-Scott, R. and McCready, L.L. (1946) Radio-Frequency Energy from the Sun. *Nature*, **157**,

- 158-159. <https://doi.org/10.1038/157158a0>
- [2] Ryle, M. and Vonberg, D.D. (1948) An Investigation of Radio-Frequency Radiation from the Sun. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, **193**, 98-120. <https://doi.org/10.1098/rspa.1948.0036>
- [3] Thompson, A.R., Moran, J.M., Swenson, G.W. (2017) *Interferometry and Synthesis in Radio Astronomy*. 3rd Edition, Springer, Cham. <https://doi.org/10.1007/978-3-319-44431-4>
- [4] Wu, X.P. (2019) *China SKA Science Report*. Science Press, Beijing.
- [5] Cornwell, T.J., Golap, K. and Bhatnagar, S. (2005) Wide Field Imaging Problems in Radio Astronomy. *Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, Philadelphia, 23 March 2005, 861-864. <https://doi.org/10.1109/ICASSP.2005.1416440>
- [6] An, T. (2019) Science Opportunities and Challenges Associated with SKA Big Data. *Science China Physics, Mechanics & Astronomy*, **62**, Article No. 989531. <https://doi.org/10.1007/s11433-018-9360-x>
- [7] Zhang, L., Xu, L. and Zhang, M. (2020) Parameterized CLEAN Deconvolution in Radio Synthesis Imaging. *Publications of the Astronomical Society of the Pacific*, **132**, Article ID: 041001. <https://doi.org/10.1088/1538-3873/ab7345>
- [8] 张利, 肖一凡, 米立功, 卢梅, 赵庆超, 王蓓, 刘祥, 张明, 谢泉. 基于 L_BFGS_B 局部极小化的自适应尺度 CLEAN 算法[J]. 贵州大学学报(自然科学版), 2021, 38(1): 38-44.
- [9] Högbom, J.A. (1974) Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astronomy and Astrophysics Supplement*, **15**, 417-426.
- [10] Cornwell, T.J. (2009) Hogbom's CLEAN Algorithm. Impact on Astronomy and Beyond. *Astronomy and Astrophysics*, **500**, 65-66. <https://doi.org/10.1051/0004-6361/200912148>
- [11] Cornwell, T.J. (2008) Multiscale CLEAN Deconvolution of Radio Synthesis Images. *IEEE Journal of Selected Topics in Signal Processing*, **2**, 793-801. <https://doi.org/10.1109/JSTSP.2008.2006388>
- [12] Bhatnagar, S. and Cornwell, T.J. (2004) Scale Sensitive Deconvolution of Interferometric Images. I. Adaptive Scale Pixel (ASP) Decomposition. *Astronomy and Astrophysics*, **426**, 747-754. <https://doi.org/10.1051/0004-6361:20040354>
- [13] Zhang, L., Zhang, M. and Liu, X. (2016) The Adaptive-Loop-Gain Adaptive-Scale CLEAN Deconvolution of Radio Interferometric Images. *Astrophysics and Space Science*, **361**, Article No. 153. <https://doi.org/10.1007/s10509-016-2746-8>
- [14] Zhang, K., Zuo, W.M., Chen, Y.J., Meng, D.Y. and Zhang, L. (2017) Beyond Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, **26**, 3142-3155. <https://doi.org/10.1109/TIP.2017.2662206>
- [15] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of Machine Learning Research*, **37**, 448-456.