

文章编号:1671-4229(2024)03-0058-07

基于星表的射电干涉阵成像评价指标研究

吴少锋, 王 锋*

(广州大学 物理与材料科学学院, 广东 广州 510006)

摘要: 为了满足新一代射电干涉阵望远镜对成像结果进行质量评价的迫切需求,文章以仿真数据为基础,通过比较在成像结果上进行源搜寻得到的星表与仿真时所使用的已知参考星表之间的差异,展开对射电干涉阵成像质量评价指标的研究。不同于常用的将两个星表进行交叉匹配并比较匹配源的诸如位置、流量密度等属性的精度,文章提出了使用度量两个元素集或者分布之间差异的推土距离(Earth Mover's Distance, EMD)来量化两个星表之间的差异,以更鲁棒的数值指标评价成像质量,为进一步研究更加鲁棒的射电干涉阵成像评价指标提供了有益的参考,也有助于射电干涉阵望远镜最终获得更高质量的科学数据。

关键词: 射电干涉阵成像; 平方公里阵列望远镜; 星表

中图分类号: P164 **文献标志码:** A

A study of the assessment metric based on catalogs for radio interferometric imaging

WU Shao-feng, WANG Feng*

(School of Physics and Materials Science, Guangzhou University, Guangzhou 510006, China)

Abstract: To meet the urgent demand for assessing the imaging quality of the next generation of radio interferometry telescopes, this study, based on simulation data, explores quality assessment metrics for radio interferometric imaging by comparing the catalogs obtained from source finding on imaging results with the known reference catalog used in simulation. Unlike the common practice of cross-matching two catalogs and comparing the accuracy of matched sources' attributes such as position and flux density, the study proposes using the Earth Mover's Distance (EMD) to quantify the difference between two sets or distributions of elements as a more robust numerical metric for assessing the quality of imaging. This research provides an important foundation for the development of more robust evaluation metrics for radio interferometric imaging, contributing to the production of higher quality scientific data by radio interferometry telescopes.

Key words: radio interferometric imaging; Square Kilometre Array telescope; catalog

0 引 言

随着央斯基甚大阵(Jansky Very Large Array,

JVLA)^[1-2]、澳大利亚平方公里阵列探路者(Australian Square Kilometre Array Pathfinder, ASKAP)^[3-4]、低频阵列(LOW Frequency ARray, LOFAR)^[5]、默奇森宽视场阵列(Murchison Widefield Array, MWA)^[6]

收稿日期: 2023-03-04; 修回日期: 2023-04-28

基金项目: 国家 SKA 专项资助项目(2020SKA011030)

作者简介: 吴少锋(1998—),男,硕士研究生. E-mail: wushaofeng@cmlab.net

*通信作者. E-mail: fengwang@gzhu.edu.cn

引文格式: 吴少锋, 王锋. 基于星表的射电干涉阵成像评价指标研究[J]. 广州大学学报(自然科学版), 2024, 23(3): 58-64.

等新一代大型射电干涉阵望远镜的建成,人类得以在较高的灵敏度、极高的分辨率下对射电天空进行观测。尤其是国际上正在建造的最大综合孔径射电望远镜——平方公里阵列(Square Kilometre Array, SKA)望远镜^[7-8],其具备极高的灵敏度(较JVLA灵敏度提高50倍、搜索速度提高10 000倍)、出色的空间分辨率,并能以纳秒级的采样精度获取细致的时间结构。由大视场、多波束、高动态范围、高分辨率和大数据等一系列最新的射电天文技术催生的SKA,为人类深入了解宇宙的形成和演化提供了前所未有的机遇。然而,随着新一代大型射电干涉阵望远镜灵敏度、时间和空间分辨率、视场的巨大提升,数据处理管线和对射电天文学至关重要算法的开发面临着巨大的挑战,尤其是在处理超大规模数据方面。

在新一代大型射电干涉阵望远镜的数据处理框架中,以宽频带进行大天区成像的连续成像管线是其中的关键部分^[9]。在大多数情况下,干涉阵列望远镜向国际天文学社区提供的标准数据产品为成像管线输出的一个图像立方体,包含两个空间维度和一个光谱频率维度。但对于新一代大型射电干涉阵望远镜来说,以SKA为例,其典型的成像结果将会是每个空间维度上有 2^{15} 像素和光谱频率维度包含高达 2^{16} 个通道,单个数据产品的体积就将达到PB级,这无疑会对数据的传输、储存和进一步处理带来非常大的挑战。为了减小成像数据产品的体积,SKA的科学数据处理器(Science Data Processor, SDP)将以成像结果——复原图像(restored image)和残差图像(residual image)的前三个泰勒系数图像作为提供给国际天文界的标准观测数据产品。泰勒系数图像是根据天体射电辐射机制的连续频率行为^[10],将图像立方体的频率轴在参考频率 ν_0 处进行泰勒展开获得的^[11]。参考频率 ν_0 处的天空亮度分布图像 I_{ν_0} 与其前三个泰勒系数图像 I_0 、 I_1 和 I_2 的对应关系为^[11]

$$I_0 = I_{\nu_0}, \quad (1)$$

$$I_1 = I_{\alpha} I_{\nu_0}, \quad (2)$$

$$I_2 = (I_{\alpha}(\alpha - 1)/2 + I_{\beta}) I_{\nu_0}, \quad (3)$$

其中, $I_{\alpha} = I_1/I_0$ 为光谱指数(spectral index)图; $I_{\beta} = (I_2/I_0) - I_{\alpha}(I_{\alpha} - 1)/2$ 为光谱曲率(spectral

curvature)图。光谱指数描述了天空亮度分布如何随频率变化,因此,通过这3个泰勒系数图像,可以还原得到包含天空亮度分布信息的图像立方体。这些以泰勒系数图像为格式的成像结果是科研人员进一步科学研究的基础。也就是说,连续成像的科学质量以及准确性直接关系着相关的重大科学研究目标能否实现。因此,射电干涉成像迫切需要一种支持泰勒展开图像的鲁棒成像质量评价方法,以确保成像的科学质量和准确性,并且该成像质量评价方法也将指导着更加精确的射电干涉成像技术的研究。

目前,Lyu等^[12]借助射电天文模拟、校准和成像库(Radio Astronomy Simulation, Calibration, and Imaging Library, RASCIL^①),并基于SKA第一阶段的低频阵列(SK phase 1 Low frequency array, SKA1-LOW)仿真数据,提出了面向SKA的射电干涉成像质量评价方法^[12]。该成像质量评价方法以连续成像管线输出的泰勒系数图像为数据基础,通过分析这些图像数据对成像质量进行评估。在这个成像质量评价方法中,关键部分为使用源搜寻软件在复原图像(restored image)的0阶泰勒系数图像上进行源搜寻并生成星表。源搜寻得到的星表在本文中被称为测试星表,而仿真时输入的已知星表被称为参考星表。该成像质量评价方法将测试星表和参考星表通过位置间隔阈值方式进行交叉匹配,然后对匹配源之间的位置、流量密度和光谱指数信息进行比较,生成一系列关于源的各个属性精度的图像指标,以评估连续成像管线的质量,如位置偏差、流量密度比率等指标。虽然这些指标都能有效地对连续成像管线的质量进行评价,但是需要维护人员或科学家进行一定程度的直接干预才能对图像形式的指标进行解读。对于具有超大规模数据量的新一代大型射电望远镜来说,尤其是SKA,将是一个巨大的挑战。因此,本研究希望提出一个数值形式并且鲁棒的指标,以对射电干涉成像的质量进行评估。本研究基于SKA1-LOW全阵列仿真数据,应用衡量两个元素集或者分布之间差异的度量指标——推土距离(Earth Mover's Distance, EMD)^[13-14]来量化测试星表和参考星表之间的差异,为提出更鲁棒的射

① <https://gitlab.com/ska-telescope/external/rascil-main>。

电干涉成像评价指标奠定了重要的研究基础。

1 星表差异量化指标:EMD

1.1 EMD

EMD 是一种衡量两个元素集或分布之间差异的度量,它基于最小化将一个分布转换为另一个分布所需的“工作量”或成本。这一概念最早由 Peleg 等^[15]在处理视觉问题时提出,并迅速被应用于图像检索等多个领域。EMD 的直观理解:给定两个分布,一个可以被视为空间中分布的土壤堆,另一个则被视为同一空间中的洞穴集合。理论上,总存在足够的土壤来填满所有洞穴至其最大容量。如果需要,可以通过调整所谓的“土壤”和“洞穴”的定义来实现这一点。因此,EMD 衡量的是填满洞穴所需的最小工作量^[14]。

EMD 的计算方法是基于解决著名的运输问题^[16],也就是 Hitchcock 在 1941 年提出的 Monge-Kantorovich 问题。这一问题可以追溯到 1781 年,最初由 Monge 提出^[17]。简单来说,如果有多个供应商,每个供应商都有一定数量的货物需要供应给多个消费者,而每个消费者都有一定的接收容量。对于每对供应商和消费者,运输单个货物单位的成本是已知的。运输问题的目标是找到一种最经济的货物流动方案,从供应商处将货物运送至消费者处,以满足所有消费者的需求。

在 EMD 的框架下,通过将一个数据集(或称为 signature)定义为供应商,另一个数据集定义为消费者,并且设定供应商-消费者对之间的成本等于两个 signature 中元素之间的地面距离(ground distance),signature 匹配自然而然被建模为一个运输问题。这里的 signature 指的是一组特征聚类,每个聚类由其均值(或模式)及相应的权重表示,其中,均值是一个 d 维向量,描述了该聚类的 d 个特征;权重则反映了该聚类在整体分布中的比重。而地面距离(ground distance)是衡量两个 signature 中元素之间差异的度量,通常基于它们特征的欧几里得距离(Euclid distance)或其他适当的距离计算方法。也就是说,EMD 度量了将一个 signature 转换为另一个 signature 所需的最小“工作量”^[14]。

EMD 的计算可以形式化为以下线性规划问题:设 $P = \{(p_1, w_{p_1}), \dots, (p_m, w_{p_m})\}$ 为具有 m 个

聚类的第一个 signature,其中, p_i 为聚类的代表(representative), w_{p_i} 是该聚类的权重; $Q = \{(q_1, w_{q_1}), \dots, (q_n, w_{q_n})\}$ 为包含 n 个聚类的第二个 signature; $D_{ij} = [d_{ij}]$ 是聚类 p_i 和 q_j 之间的地面距离。研究希望找到将 P 转换到 Q 的最小总流(total flow) $F = [f_{ij}]$, f_{ij} 为 p_i 与 q_j 之间的流:

$$\text{WORK}(P, Q, F) = \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}, \quad (4)$$

有以下约束条件:

$$f_{ij} \geq 0, 1 \leq i \leq m, 1 \leq j \leq n; \quad (5)$$

$$\sum_{j=1}^n f_{ij} \leq w_{p_i}, 1 \leq i \leq m; \quad (6)$$

$$\sum_{i=1}^m f_{ij} \leq w_{q_j}, 1 \leq j \leq n; \quad (7)$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min\left(\sum_{i=1}^m w_{p_i}, \sum_{j=1}^n w_{q_j}\right); \quad (8)$$

其中,约束条件式(5)允许从 P 向 Q 移动“物资”,而不能反向移动;约束条件式(6)限制了 P 中的聚类可以发送的物资量不超过其权重;约束条件式(7)限制了 Q 中的聚类接收的物资量不超过其权重;而约束条件式(8)要求移动尽可能多的物资。一旦运输问题得到解决,并且找到了最优的总流 F ,则推土距离被定义为由总流归一化的工作量:

$$\text{EMD}(P, Q) = \frac{\sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}}, \quad (9)$$

其中,分母是一个归一化因子,是较小 signature 的总权重,这是因为约束条件式(8)的要求。当两个 signature 具有不同的总权重时,则需要这个因子进行归一化,以避免偏向于较小的 signature。一般情况下,地面距离 d_{ij} 可以是任意距离,需要根据具体的问题来选择。因此,EMD 自然地将单个元素之间的距离概念扩展到元素集或分布之间的距离,并且以一种自然的方式允许部分匹配^[14]。

1.2 星表间差异的 EMD 量化

由 Lyu 等^[12]提出的成像质量评估方法基于源搜寻结果,其评价指标主要基于测试星表与参考星表之间匹配源的各个属性差异。然而,这种评价只考虑了能够在测试星表中与参考星表匹配的源,忽略了无法匹配的源。实践中,使用不同的成像算法或参数对同一组可见度数据进行成像,并利用相同的源搜寻软件和参数进行源搜寻时,通常会得到不同的搜寻结果。这种差异不仅体现在

匹配源的流量密度、位置和光谱指数的准确性不同上,还体现在产生的错误检测源上,这也是评估成像质量时不可忽视的方面。因此,考虑测试星表中所有源的全面分析将提供更鲁棒的成像质量评估结果。

EMD 具备衡量元素集合或分布之间差异的能力,且适用于部分匹配的情况,为评估测试星表与参考星表之间的差异提供了一个新角度。基于 EMD 的原理和其应用范围可以发现,在基于星表的成像质量评价方法中使用 EMD 来衡量两个星表之间的差异具有重要的潜力,它可以作为一个更鲁棒的数值指标来反映成像质量。EMD 能更全面地分析测试星表与参考星表的整体差异,从而帮助更加全面和精确的成像质量评价。

本研究专注于一个仅包含点源的天空模型,此时,参考星表和测试星表都处于简化的情况:只包含点源,每个点源由赤经、赤纬、流量密度和光谱指数 4 个主要属性描述。因此,可以把星表视作一个由具有这 4 种特征的元素组成的集合。在射电天文源搜寻中,星表的完整性和可靠性通常与点源的流量密度正相关,表明在比较测试星表和参考星表时,流量密度是一个关键属性。因此,可以将点源的流量密度作为其在星表中的权重,使得每个星表在 EMD 框架下可以被当做一个独特的 signature。在这种框架下,每个点源构成一个以流量密度为权重的聚类,特征向量由其赤经、赤纬和光谱指数组成。

当将测试星表和参考星表视作 EMD 框架下的 signature 进行计算时,必须对数据进行适当的定义以满足 EMD 的 4 项约束条件。由于源搜索软件可能无法召回所有真实源,使得测试星表中的源成为参考星表中源的一个子集。此外,射电干涉阵数据受主波束效应影响,可能会导致测试星表中的源流量密度普遍低于其真实值。基于这些考虑,本研究将参考星表作为供应商(signature 1),测试星表作为消费者(signature 2)。此时计算得到的 EMD 值代表了从参考星表(供应商)到测试星表(消费者)的转移“成本”,该成本量化了两个星表之间的差异。在这个框架下,更小的 EMD 值意味着测试星表与参考星表之间的差异更小,

指示出更高的成像质量,为射电干涉阵成像的质量评估提供了一种更鲁棒的指标。

2 EMD 在仿真数据上的实测

本研究基于 SKA1-LOW 全规模仿真数据,对应用 EMD 量化测试星表和参考星表的差异以评估射电干涉成像质量进行了实测。本测试使用了仿真数据,其优势在于能够获得参考星表的所有真实信息。

2.1 SKA1-LOW 全规模仿真数据

SKA1-LOW 全阵列将由 131 072 个对数偶极子天线组成,这些天线分布在 512 个站点上,每个站点中包含 256 个天线。每个站点等效为口径 38 m 的孔径望远镜,大约 50% 的站点将被布置在直径为 1 km 的核心区域内,其余的站点则分布在 3 个螺旋线上,每个螺旋线上有 6 个站点的集群。

本研究使用 GPU 加速的 SKA 仿真软件 OSKAR^① 进行模拟观测,生成 SKA1-LOW 全阵列的仿真可见度数据。仿真输入的天空模型为真实星表和随机源的组合,选择使用真实星表是因为其中的源的各个属性都来自于对真实的天空的观测,更接近实际的观测情况。本研究使用的真实星表为 GLEAM (GaLactic and Extragalactic All-sky Murchison Widefield Array) 巡天项目中的星表^[18]。GLEAM 巡天为 MWA 的一个重要巡天项目。天空模型除了真实的 GLEAM 星表之外,还包括 OSKAR 生成的均匀分布在全天的 8 000 000 个随机点源。这些随机点源的流量密度范围在 $10^{-4} \sim 10^{-1}$ Jy 之间,遵循幂指数为 1 的幂律分布。仿真过程中,参考了 SKA 的主要科学目标——探测宇宙再电离(Epoch of Reionization, EoR)信号的深度成像候选场^[19],设定 $(347.5^\circ, -8.1^\circ)$ 作为模拟观测的相位中心,视场为 20 平方度。在实际的射电干涉测量中,干涉仪主波束外的天空也会进入到测量数据中。因此,本研究对天空模型进行了筛选,选择该相位中心 3.5° 内的 GLEAM 星表中的源和随机弱源作为仿真的天空模型。

仿真的总观测时长为 14 400 s,共有 240 个时间采样点,每个采样的积分时间为 0.9 s。为了后续

① <https://github.com/OxfordSKA/OSKAR>。

连续成像处理中能够有足够的通道数进行 Taylor 展开,同时为平衡计算资源,模拟观测的通道数设置为 96 个。模拟观测的起始频率为 72 MHz,每个通道的带宽为 10 kHz,通道之间的间隔为 200 kHz。使用 OSKAR 进行观测的主要参数汇总为表 1。

表 1 OSKAR 仿真中的主要参数

参数名称	参数值
天空模型	GLEAM + random
流量密度范围/Jy	$10^{-4} \sim 2.93$
望远镜模型	SKA1 - LOW full array
相位中心/degree	(347.5, -8.1)
起始频率/MHz	72
通道数/个	96
通道间隔/kHz	200
通道带宽/kHz	10
观测时长/s	14 400
时间采样个数/个	240
单次积分时间/s	0.9

在实际观测中,射电干涉阵望远镜需要根据不同的观测模式设定相应的参数。考虑到不同观测参数可能引入的热噪声水平变化,本研究在 OSKAR 生成的仿真可见度数据中添加了不同级别的热噪声,范围为 $1 \sim 6\sigma$, 1σ 对应 10^{-5} Jy 的热噪声均方根起伏(Root Mean Square, RMS)水平。

本研究使用 RASCIL 中的连续成像管线对 OSKAR 生成的 SKA1-LOW 全阵列仿真可见度数据进行连续成像。为了将连续成像的结果保存为低内存占用的泰勒系数图像,使用多尺度多频率(Multi-Scale Multi-Frequency Synthesis, MSMFS)退卷积算法^[11]对 SKA1-LOW 全阵列仿真可见度数据进行退卷积,并设定该管线输出的泰勒系数图像为 3 个。因此,连续成像的结果为 SKA1-LOW 全阵列仿真可见度数据的复原图像和残差图像在平均频率 81.6 MHz 处的前三个泰勒系数图像。关于成像视场大小的设置,考虑到所选择的 EoR 深度成像候选场的大小 20 平方度,所以成像视场的赤经、赤纬轴尺度被设置为 4.5° 。成像分辨率需要根据图像的大小进行设置,以图像大小为 $16\,384 \times 16\,384$ 像素的仿真为例子,图像的每个像素的大小应该设置为 4.7937×10^{-6} rad。成像权重设置为 Robust(-2.0)权。使用 RASCIL 中的连续成像管线对 OSKAR 生成的 SKA1-LOW 全阵列仿真可见度数据进行连续成像的主要参数汇总为表 2。

表 2 RASCIL 连续成像管线的主要参数

参数名称	参数值
洁化算法	MSMFS
Major 循环数/个	8
Minor 循环数/个	10 000
洁化阈值/Jy	10^{-4}
洁化切片数/个	4
成像视场/degree ²	4.5×4.5
成像尺寸/pixel ²	$16\,384 \times 16\,384$
成像分辨率/rad	4.7937×10^{-6}
成像权重	Robust(-2.0)
输入通道数/个	96
MFS 通道数/个	6
输出格式	Taylor
展开阶数/个	3

2.2 基于仿真数据对 EMD 指标的测试

本研究基于仿真生成的不同热噪声水平的 SKA1-LOW 全阵列仿真数据,应用 EMD 来量化测试星表与参考星表之间的差异,从而评价射电干涉阵成像的质量。

为了更全面地测试 EMD 指标在评价射电干涉成像质量中的表现,本研究使用了射电天文学中广泛使用的 4 种源搜寻软件:①GLEAM 巡天中的源搜寻软件 Aegean^[20-21];②为 SCORPIO(Stellar Continuum Originating from Radio Physics in Our galaxy)项目和下一代射电巡天而设计的 Caesar^[22-24];③为满足 LOFAR 的需要而开发的 PyBDSF^[25];④WALLABY(Widefield ASKAP L-band Legacy All-sky Blind Survey)项目应用的 SoFIA^[26-28],分别对不同热噪声水平的成像结果进行源搜寻,以生成相应的测试星表。

在本研究的仿真中,对不同热噪声级别所使用的连续成像参数是一模一样的,是为了确保成像结果间的差异仅由热噪声水平的变化引起。这种方法可以预估热噪声水平差异对成像结果质量的影响。当使用相同的源搜索软件以相同的参数对这些成像结果进行源搜索时,生成的测试星表与参考星表之间的差异预期会随着热噪声的增强而增大,即同种源搜寻软件生成的测试星表与参考星表之间的 EMD 值预期会随着热噪声的增强而增大。

EMD 值的计算基于 Python 语言,借助开源计算机视觉库(Open Source Computer Vision Library, OpenCV³) Python 接口中的 cv2.EMD 函数来实现测

试星表和参考星表之间的 EMD 计算。cv2. EMD 函数需要输入两个 signature 以及所使用的地面距离类型。在本研究使用的天空模型中,每个点源以流量密度作为权重,其特征包含天球坐标赤经和赤纬以及光谱指数。为了可以使用欧几里得距离作为计算星表间 EMD 时的地面距离类型,本研究在实现时忽略了源的光谱指数特征。使用这种方式,分别计算了参考星表和 Aegean、Caesar、PyBDSF 和 SoFIA 在不同热噪声水平数据的成像结果上进行源搜寻生成的星表,即测试星表之间的 EMD,结果如图 1 所示。

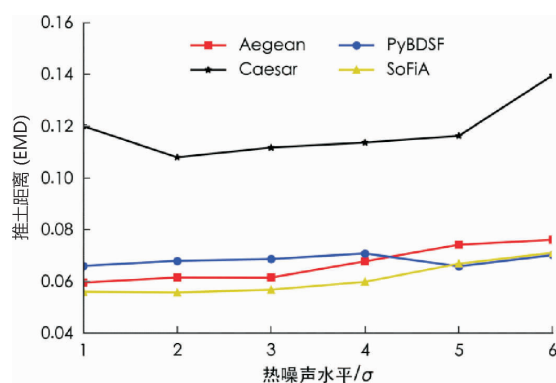


图 1 不同热噪声水平下各个源搜寻软件星表和参考星表之间的 EMD 值

Fig. 1 EMDs between the reference catalogs with catalogs generated by source finders under different thermal noise levels

从图 1 中可以看到,在不同热噪声水平的仿真数据上,这 4 种源搜寻软件生成的测试星表与参考星表之间的 EMD 值总体上呈现出随着热噪声的增强而增大的趋势,尤其是 Aegean 和 SoFIA。因此,EMD 能够在一定程度上反映出数据中不同强度的热噪声,可以作为有效的成像质量评价指标。

3 结论和展望

本研究以 SKA1-LOW 全阵列的仿真数据为基础,初步探究了在射电干涉阵成像质量评价中应用 EMD 来量化测试星表与参考星表之间的差异。研究表明,EMD 能够在一定程度上有效反映数据中不同级别的热噪声。同时,EMD 作为一个数值形式的指标,便于新一代大型射电干涉阵望远镜实现自动化的成像质量评价。本研究为提出更鲁棒的射电干涉阵成像星表评价指标奠定了重要基础,有助于新一代大型射电干涉阵列望远镜为科学研究提供更高质量的数据。

为了更加准确地使用 EMD 量化测试星表和参考星表之间的差异,下一步工作将对包含光谱指数特征的地面距离进行探究,使得 EMD 成为一个更加鲁棒的基于星表的射电干涉阵成像评价指标。

参考文献:

- [1] Napier P J, Thompson A R, Ekers R D. The very large array: Design and performance of a modern synthesis radio telescope [J]. Proceedings of the IEEE, 1983, 71(11): 1295-1320.
- [2] Thompson A R, Clark B G, Wade C M, et al. The very large array [J]. The Astrophysical Journal Supplement Series, 1980, 44(2): 151-167.
- [3] Johnston S, Taylor R, Bailes M, et al. Science with ASKAP: The Australian square-kilometre-array pathfinder [J]. Experimental Astronomy, 2008, 22(3): 151-273.
- [4] Norris R P, Hopkins A M, Afonso J, et al. EMU: Evolutionary map of the universe [J]. Publications of the Astronomical Society of Australia, 2011, 28(3): 215-248.
- [5] Van Haarlem M P. LOFAR: The low-frequency array [J]. EAS Publication Series, 2005, 15: 431-444.
- [6] Tingay S J, Goeke R, Bowman J D, et al. The Murchison widefield array: The square kilometre array precursor at low radio frequencies [J]. Publications of the Astronomical Society of Australia, 2013, 30: e007.
- [7] Dewdney P E, Hall P J, Schilizzi R T, et al. The square kilometre array [J]. Proceedings of the IEEE, 2009, 97(8): 1482-1496.
- [8] Braun R. The Square Kilometer Array interferometer [EB/OL]. (1995-12-11) [2024-03-01]. <https://doi.org/10.48550/arXiv.astro-ph/9512060>.
- [9] Dewdney P. SKA1 design baseline description [EB/OL]. (2022-01-24) [2024-03-01]. https://www.dropbox.com/scl/fi/z75259ah3hlyrcnrdocyo/SKA-TEL-SKO-0001075-02_DesignBaselineDescription.pdf?rlkey=rku3q8w04zim47o1ywj4qccv6&dl=0.

- [10] Scaife A M M. Big telescope, big data: Towards exascale with the square kilometre array[J]. *Philosophical Transactions of the Royal Series A: Mathematical, Physical, and Engineering Sciences*, 2020, 378(2166): 20190060.
- [11] Rau U, Cornwell T J. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry [J]. *Astronomy & Astrophysics*, 2011, 532: A71.
- [12] Lyu Y H C, Hodosán G, Wang F, et al. 2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC), May 30—June 04, 2022 [C]. Piscataway: IEEE, 2022.
- [13] Rubner Y, Tomasi C, Guibas L J. Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), January 07, 1998 [C]. Piscataway: IEEE, 1998.
- [14] Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval [J]. *International Journal of Computer Vision*, 2000, 40(2): 99-121.
- [15] Peleg S, Werman M, Rom H. A unified approach to the change of resolution: Space and gray-level [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(7): 739-742.
- [16] Hitchcock F L. The distribution of a product from several sources to numerous localities [J]. *Journal of Mathematics and Physics*, 1941, 20(1/2/3/4): 224-230.
- [17] Rachev S T. The monge-kantorovich mass transference problem and its stochastic applications [J]. *Theory of Probability & its Applications*, 1985, 29(4): 647-676.
- [18] Hurley-Walker N, Callingham J R, Hancock P J, et al. GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey -I. A low-frequency extragalactic catalogue [J]. *Monthly Notices of the Royal Astronomical Society*, 2017, 464(1): 1146-1167.
- [19] Zheng Q, Wu X P, Guo Q, et al. Pre-selection of the candidate fields for deep imaging of the epoch of reionization with SKA1-low [J]. *Monthly Notices of the Royal Astronomical Society*, 2020, 499(3): 3434-3444.
- [20] Hancock P J, Murphy T, Gaensler B M, et al. Compact continuum source finding for next generation radio surveys: Compact continuum source finding [J]. *Monthly Notices of the Royal Astronomical Society*, 2012, 422(2): 1812-1824.
- [21] Hancock P J, Trott C M, Hurley-Walker, N. Source finding in the era of the SKA (precursors): Aegean 2.0 [J]. *Publications of the Astronomical Society of Australia*, 2018, 35: e011.
- [22] Umana G, Tringilio C, Franzen T M O, et al. SCORPIO: A deep survey of radio emission from the stellar life-cycle [J]. *Monthly Notices of the Royal Astronomical Society*, 2015, 454(1): 902-912.
- [23] Raggi S, Ingallinera A, Leto P, et al. Automated detection of extended sources in radio maps: Progress from the SCORPIO survey [J]. *Monthly Notices of the Royal Astronomical Society*, 2016, 460(2): 1486-1499.
- [24] Raggi S, Vitello F, Becciani U, et al. Caesar source finder: Recent developments and testing [J]. *Publications of the Astronomical Society of Australia*, 2019, 36: e037.
- [25] Mohan, N, Rafferty, D. PyBDSF: Python blob detection and source finder [CP]. *Astrophysics Source Code Library*, 2015; ascl:1502.007.
- [26] Koribalski B S, Staveley-Smith L, Westmeier T, et al. WALLABY – An SKA Pathfinder H i survey [J]. *Astrophysics and Space Science*, 2020, 365(7): 118.
- [27] Serra P, Westmeier T, Giese N, et al. SoFiA: A flexible source finder for 3D spectral line data [J]. *Monthly Notices of the Royal Astronomical Society*, 2015, 448(2): 1922-1929.
- [28] Westmeier T, Kitaeff S, Pallot D, et al. SOFIA 2 – An automated, parallel H i source finding pipeline for the WALLABY survey [J]. *Monthly Notices of the Royal Astronomical Society*, 2021, 506(3): 3962-3976.

【责任编辑: 孙向荣】