

结合基于密度的噪声应用空间聚类 and 反距离加权算法的城市环境电磁频谱地图构建

谢佳炜, 余志勇, 吕典, 刘杨秋子
(火箭军工程大学, 陕西 西安 710025)

摘 要: 针对传统电磁频谱地图构建方法感知节点分布不均匀、监测数据存在异常值等问题, 将基于密度的噪声应用空间聚类 (Density-based Spatial Clustering of Applications with Noise, DBSCAN) 算法与反距离加权 (Inverse Distance Weighting, IDW) 算法相结合, 提出一种城市环境电磁频谱地图构建方法。该算法首先通过 DBSCAN 减轻极端值的影响, 并分离核心点、边界点和噪声点, 将每个核心点的局部密度作为权重, 计算簇的加权中心点。其次, 运用 IDW 对聚类簇的加权中心点进行插值估计, 以显著减少需要进行插值的数据点数量, 从而构建精度更高的电磁频谱地图。仿真结果表明: 与 IDW 算法和反障碍距离加权算法相比, 所提算法重构得到的电磁频谱地图的平均绝对误差和归一化均方误差分别降低了 10% 和 23% 以上。

关键词: 电磁频谱地图; 反距离加权; 城市环境; 密度聚类; 空间插值

中图分类号: TN97

文献标志码: A

文章编号: 2097-2741(2024)04-093-006

Construction of Electromagnetic Spectrum Map for Urban Environment Based on DBSCAN-IDW Algorithm

XIE Jiawei, YU Zhiyong, LYU Dian, LIU Yangqiuzi
(Rocket Force University of Engineering, Xi'an 710025, Shaanxi)

ABSTRACT: Aiming at the problems of uneven distribution of sensing nodes and abnormal values in monitoring data, an urban environment electromagnetic spectrum map construction method was proposed by combining density-based spatial clustering of applications with noise (DBSCAN) and inverse distance weighting (IDW) algorithm. Firstly, DBSCAN was used to reduce the influence of extreme values, and separate the core points, boundary points and noise points. The local density of each core point was used as weight to calculate the weighted center point of the cluster. Secondly, IDW was used to interpolate the weighted center point of the cluster, by which the number of data points needed to be interpolated could be significantly reduced and the electromagnetic spectrum map would be more accurate. Simulation results showed that compared with the IDW algorithm and the VIDW algorithm, the mean absolute error and the normalized mean square error of the electromagnetic spectrum map reconstructed by the proposed algorithm can be reduced by more than 10% and 23% respectively.

收稿日期: 2024-04-07

第一作者: 谢佳炜 (1993—), 女, 硕士研究生, 主要从事电磁态势感知研究。E-mail: 986932773@qq.com

引用格式: 谢佳炜, 余志勇, 吕典, 等. 结合基于密度的噪声应用空间聚类 and 反距离加权算法的城市环境电磁频谱地图构建[J]. 火箭军工程大学学报, 2024, 38 (4): 93-97, 107.

KEYWORDS: electromagnetic spectrum map; inverse distance weighting; urban environment; density clustering; spatial interpolation

电磁频谱地图 (Electromagnetic Spectrum Map, EMSM) 是一种多维度的数据表达形式, 特定区域内呈现出不同频段的电磁波时空分布和传播特性。进一步结合地理信息系统, 还可以清晰展示电磁环境的能量和频谱资源分布情况, 为优化频谱利用和干扰管理策略提供可视化的数据支持^[1]。电磁频谱地图的生成技术或过程也称为电磁频谱测绘, 是电磁环境建模与仿真的基础, 对于确保电磁环境建模与仿真的准确性和可靠性具有重要作用。

电磁频谱地图构建方法主要分为 3 类: 空间插值法, 仅依赖传感器采集的电磁环境数据进行估计; 参数构建法, 在监测数据基础上融入发射机信息及传播模型等先验知识; 混合构建法, 则是前两者的综合应用^[2]。基于空间插值的算法因无需依赖先验知识而备受关注^[3]。该方法利用监测区域内监测站点采集的信号接收功率或接收信号强度 (Received Signal Strength, RSS) 数据, 即可获得贴近实际的构建结果, 实现高精度电磁频谱地图构建, 从而具有更广泛的应用前景^[4]。

基于空间的插值算法中, 反距离加权 (Inverse Distance Weighting, IDW) 算法通过计算预测点与已知样点之间的距离来确定权重, 从而对未知点的属性值进行估计^[5]。这种算法具有原理简单、快速、易于解释的优点, 在许多领域得到了广泛应用。吴昱江等^[6-7]针对传统 IDW 算法受限于插值邻域需人工选择等局限性, 设计了一种基于 Voronoi 图反障碍距离加权 (Voronoi-based Inverse Obstacle Distance Weighted, VIODW) 算法, 显著提升了复杂场景下电磁频谱地图的构建性能。赵亚凯等^[8]对反距离加权插值算法进行了深入研究, 通过调整幂函数、相邻点数、搜索形状和搜索扇区等参数, 成功构建数字高程模型 (Digital Elevation Model, DEM)。Yapar 等^[9]对 IDW 算法进行了总结, 指出其存在的缺陷, 并提出了一种自适应的距离衰减参数策略。这种策略能够更准确地模拟研究区域内距离衰减关系的变化, 取得了更好的插值结果。尽管关于 IDW 算法的理论推导和改进研究层出不穷, 但多数研究仍然聚焦于幂参数的选取和优化以及基于搜索半径的选择。这些改进在一定程度上提升了插值效果,

但往往忽视了异常值和感知节点分布不均所带来的“边界效应”对插值结果的潜在影响。

本文针对传统 IDW 算法在构建电磁频谱地图时存在需要反复全局性使用所有监测站点数据、对异常值敏感且效率不高等问题, 提出了一种基于密度的噪声应用空间聚类 (Density-based Spatial Clustering of Applications with Noise, DBSCAN) 与 IDW 算法相结合的电磁频谱地图生成技术。

1 IDW 算法原理

IDW 是一种全局性、确定性的空间插值方法。在形式上, IDW 算法可利用空间中 N 个已知观测点位置 (x_i, y_i) 处的观测值 z_i 来估计未知点 (x, y) 处的未知值 z , 即

$$z(x, y) = \frac{\sum_{i=1}^N w_i z_i}{\sum_{i=1}^N w_i}$$

式中: w_i 为第 i 个点的权重, 即已知点的函数值进行加权平均后可得待插值位置的函数值。 w_i 与待估计点 $z(x, y)$ 之间的欧几里得距离成反比, 即

$$w_i = \frac{1}{d_i^\alpha}$$

式中: d_i 为已知点和待插值点的距离; α 为可调参数, 用于控制距离权重的衰减速率, 一般取 2。

2 DBSCAN-IDW 插值算法

由于传统 IDW 算法构建电磁频谱地图时存在需要反复全局性使用所有监测站点数据、对异常值敏感且效率不高等问题, 本文提出通过使用密度聚类算法 (如 DBSCAN) 时^[10-11], 算法会根据指定的参数 (邻域半径 ε 和最小样本数 MinPts) 识别出核心点, 聚类完成后对簇内核心点进行加权平均, 选取这些中心点作为插值计算中的已知点, 将聚类得到的中心点作为“控制”点, 进而运用 IDW 方法插值得到目标区域内待测点的信号强度值, 以提升城市环境下电磁频谱地图的构建精度^[12-13]。

2.1 算法流程

本文所提出的 DBSCAN-IDW 算法主要包括通过使用密度聚类算法实现高低密度划分（核心点、边界点、噪声点）和 IDW 算法插值 2 个核心部分，分别完成待插值点的邻域划分和基于划分的结果来计算待插值点值 2 个任务，最终得到高精度的电磁频谱地图。

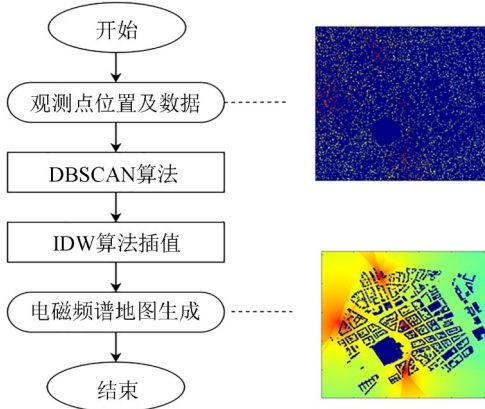


图 1 DBSCAN-IDW 算法流程

Fig. 1 Procedure of DBSCAN-IDW algorithm

2.2 算法步骤

根据各个测量点的位置分布及其对应的信号强度值，进行密度聚类，以实现相似划分。假定目标区域 S 随机初始的 2 个关键参数——邻域半径 ε 和最小样本数 MinPts 已知，在数据集上运行 DBSCAN 算法，将数据点划分为不同的簇。

$$p \in NE_{ps}(q)$$

$$|NE_{ps}(q)| \geq \text{MinPts}$$

式中： p 为核心点； q 为当前点； NE_{ps} 为邻域半径 ε ；最大值 $NE_{ps}(q)$ 为样本集的点的数量； $|\cdot|$ 表示集合的元素数量。

随机选择一个未被访问过的点 q 作为当前点，如果当前点的邻域半径 ε 内的点数量大于等于 MinPts ，则当前点 q 为核心点。否则，当前点为边界点或噪声点。从核心点开始，将其标记为簇的一部分，并递归地访问其邻域半径内的所有点。如果某个点的 ε 邻域内的点数量大于等于 MinPts ，则该点为核心点，并将其标记为已访问。重复此步骤直到所有点都被访问，即得所有簇的结果。

每个簇包含一组密度可达的点，共可分 B 簇（即 B 个聚类中心）， $B = (1, 2, \dots, K)$ 。如果簇 B_k 包含 n 个核心点（边界点和噪声点忽略不计），并且每个核心点在二维空间中的位置由坐标 (x_{ki}, y_{ki})

表示，使用每个核心点的局部密度作为权重，计算簇的加权中心点。而簇中的每个核心点的局部密度是基于邻居数量的局部密度，使用 DBSCAN 中的邻域半径 ε ，统计每个点 ε 半径内的邻居数量（包括点本身）作为该点的局部密度。加权中心是簇内所有核心点的位置与其对应权重的加权平均，即得计算加权中心的坐标和信号强度值。

$$x_{kc} = \frac{\sum_{i=1}^n \rho_i \cdot x_{ki}}{\sum_{i=1}^n \rho_i}$$

式中： x_{kc} 是加权中心的横坐标； n 是簇中监测点的数量； ρ_i 是第 i 个点的局部密度； x_{ki} 是第 i 个点的横坐标。同理，可得加权中心的纵坐标 y_{kc} 。

这样，簇 B_k 的加权中心坐标就是 (x_{kc}, y_{kc}) 。同理可得聚类密度加权中心的电场强度值。对于未知待测量点 (x_0, y_0) ，计算其与各个簇内中心点欧式距离为

$$d(0, kc) = \sqrt{(x_0 - x_{kc})^2 + (y_0 - y_{kc})^2}$$

选择聚类簇的权重中心点作为已知点，利用计算出的距离和已知中心点的坐标来计算未知点的预估值。IDW 的基本公式是

$$E(x) = \frac{\sum_{k=1}^B \left(\frac{1}{d(x_0, x_{kc})^\alpha} E(x_k) \right)}{\sum_{k=1}^B \left(\frac{1}{d(x_0, x_{kc})^\alpha} \right)}$$

式中： $E(x)$ 是插值点的预估值； $E(x_k)$ 是第 k 个已知中心点的值； $d(x_0, x_{kc})$ 是插值点与已知中心点间的距离。

2.3 算法性能评估指标

本文采用平均绝对误差 E_{MAE} 和归一化均方误差 E_{NMSE} 作为评估电磁频谱地图构建精度的指标。 E_{MAE} 表示真实信号强度 $p_n(s_n)$ 和估计信号强度 $\hat{p}_n(s_n)$ 的平均绝对值误差， E_{NMSE} 表示真实信号强度和算法估计信号强度的归一化均方误差，数学表达式分别为

$$E_{MAE} = \frac{1}{N} \sum_{n=1}^N |p_n(s_n) - \hat{p}(s_n)|$$

$$E_{NMSE} = \frac{1}{N} \sum_{n=1}^N \frac{(p_n(s_n) - \hat{p}_n(s_n))^2}{(p_n(s_n) - \bar{p}_n(s_n))^2}$$

式中： N 表示电磁频谱的聚类簇个数； $\bar{p}_n(s_n)$ 表示真实信号强度平均值。

3 城市环境建模与仿真实验

3.1 仿真环境建模

为验证本文提出的 DBSCAN-IDW 算法在特定环境下对电磁频谱的预测性能, 首先使用电磁波传播和无线网络规划软件导入某城市环境模型, 并使用 Altair WinProp 软件中优势路径射线追踪传播模型, 获取理论电磁频谱地图以模拟算法的应用环境。在图 2 环境内模拟了一个面积为 $1.8 \text{ km} \times 1.2 \text{ km}$ 的城市二维地形目标区域, 并离散为 $n=100$ 的网格大小。在这个模型中, 设定了一个树木带 (图中绿色部分, 高度为 $10 \sim 30 \text{ m}$), 以及建筑物 (图中灰色结构, 高度为 $50 \sim 110 \text{ m}$), 衰减属性为 0.25 dB/m 。



图 2 某城市二维平面地形图

Fig. 2 2D topographic map of a city

该环境模型的三维地形如图 3 所示。在环境模型建筑物最高处放置了一个高度为 40 m 、频率为 2.5 GHz 、功率为 40 W 的全向发射天线 (图中蓝色部分)。设置了路径损耗指数为 $2.7 \sim 3.5$, 其电磁频谱地图空间分辨率为 10 m 。

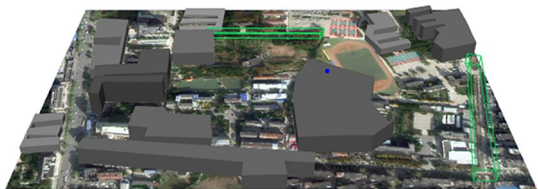


图 3 某城市三维地形图

Fig. 3 3D topographic map of a city

在建模过程中, 采用优势路径射线追踪传播模型获取理论电磁频谱地图。为不失一般性和便于计算, 场景设计简化, 即同一目标区域特定时间段在单个辐射源场景下构建城市电磁频谱地图,

进而仿真得到理论电磁频谱地图, 如图 4 所示。

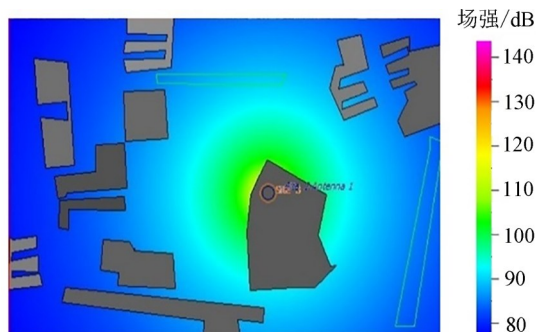


图 4 某城市环境下理论电磁频谱地图

Fig. 4 Theoretical electromagnetic spectrum map in an urban environment

3.2 仿真实验结果分析

在得到理论电磁频谱地图的基础上, 使用拉丁超立方抽样 (Latin Hypercube Sampling, LHS) 方法随机选取不同的监测站点。通过获取这些电磁监测站点的电场强度值, 随机抽取 $2\,000$ 个监测点的数据, 分别使用 Kriging、IDW、VIDW 和本文提出的 DBSCAN-IDW 算法对目标区域进行插值估计, 进而生成城市环境下的电磁频谱地图。不同算法的 E_{MAE} 和 E_{NMSE} 如图 5 和图 6 所示。生成的地图如图 7 所示。

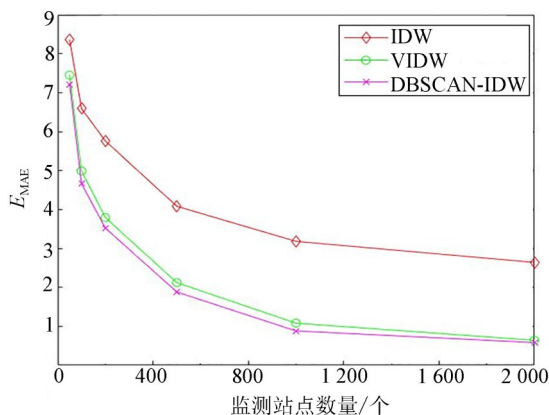


图 5 3 种算法的平均绝对误差性能对比曲线

Fig. 5 Curves of MAE error performance of the three algorithms

由图 5 可看出: DBSCAN-IDW 算法在城市环境电磁频谱预测中具有较高的准确性和适用性, 且在不同数量的监测站点下均表现出较好的预测性能; 随着监测站点数量增加, 其平均绝对误差呈下降趋势, 当监测点的数量达到 $2\,000$ 时, 各个算法平均绝对误差分别为 $2.773\,6$ 、 $0.683\,1$ 和 $0.633\,1$; DBSCAN-IDW 算法的 E_{MAE} 比传统的

IDW 算法和 VIDW 算法分别下降了约 12% 和 10%。

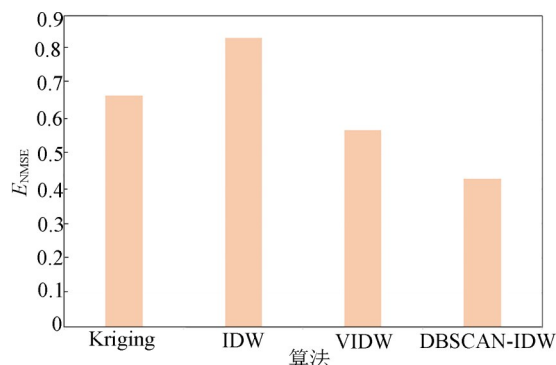


图 6 不同算法的归一化均方误差性能对比

Fig. 6 Performance comparison of normalized mean square errors of different algorithms

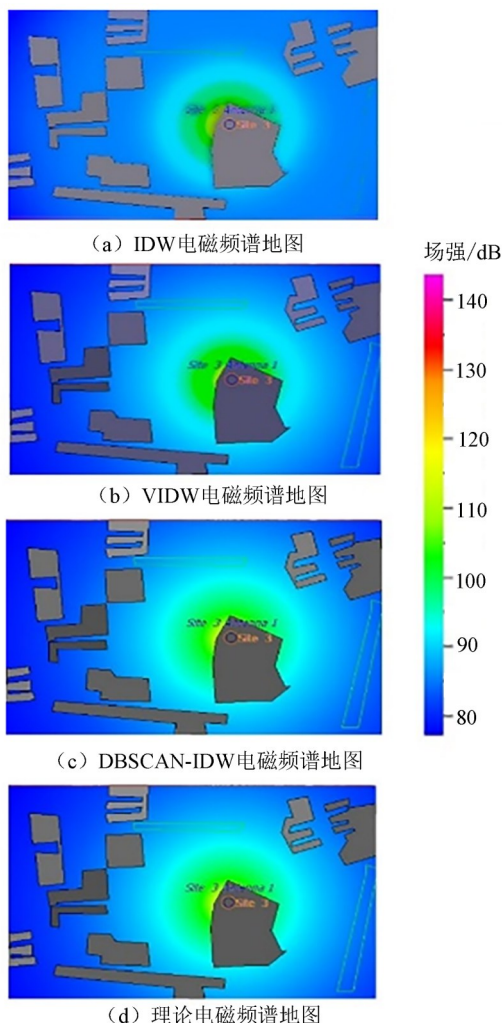


图 7 3 种算法生成的电磁频谱地图

Fig. 7 Electromagnetic spectrum maps generated by three algorithms

从图 6 可以看出, DBSCAN-IDW 算法的 E_{NMSE} 值比 Kriging、VIDW、IDW 算法降低 23%~

37%, 表明其电磁频谱地图构建精度明显提升。

4 结论

本文针对传统 IDW 算法在构建电磁频谱地图时存在需要反复全局性使用所有监测站点数据、对异常值敏感且效率不高等问题, 提出了一种 DBSCAN-IDW 算法的电磁频谱地图生成技术。通过对城市环境进行建模仿真, 验证了该算法在城市环境下能够构建更高精度的电磁频谱地图。但是本文的密度聚类考虑不够全面, 只考虑聚类簇核心点, 忽略边界点和噪声点会影响局部构建结果。接下来会进一步研究考虑聚类边界点, 使频谱地图构建更加精准和贴合实际。

参考文献:

- [1] 夏海洋, 查淞, 黄纪军, 等. 电磁频谱地图构建方法研究综述及展望[J]. 电波科学学报, 2020, 35(4): 445-456.
- [2] 胡炜林, 朱然刚, 彭闯, 等. 电磁环境态势生成与认知技术研究现状与展望[J]. 探测与控制学报, 2022, 44(6): 8-17.
- [3] 查淞, 夏海洋, 黄纪军, 等. 基于监测数据的电磁频谱地图构建与验证[J]. 国防科技大学学报, 2023, 45(3): 171-178.
- [4] 李赞, 韩宇. 一种稀疏采集条件下频谱地图构建方法: CN115988513A[P]. 2023-04-18.
- [5] 方胜良, 胡豪杰, 储飞黄, 等. 一种基于改进反距离插值法的电磁频谱地图实现方法: CN115100376A[P]. 2022-09-23.
- [6] 吴昱江, 丁浩, 王元贺, 等. 基于 VIDW 算法的城市环境电磁频谱地图构建[C]//第十八届全国电波传播年会论文集. 北京: 中国电子学会电波传播分会, 2023: 658-661.
- [7] 陶诗飞, 吴昱江, 罗佳, 等. 基于反障碍距离加权的

- 复杂场景电磁频谱地图构建方法[J]. 电子与信息学报, 2024, 46(7): 1-9.
- TAO S F, WU Y J, LUO J, et al. Radio environment map construction method for complex scenes based on inverse obstacle distance weighted [J]. Journal of Electronics & Information Technology, 2024, 46(7): 1-9.
- [8] 赵亚凯, 邓青春. 反距离加权插值参数对细沟 DEM 精度的影响[J]. 西华师范大学学报(自然科学版), 2023, 44(5): 496-504.
- ZHAO Y K, DENG Q C. Impact of inverse distance weighted interpolation parameters on rill DEM accuracy[J]. Journal of China West Normal University(Natural Sciences), 2023, 44(5): 496-504.
- [9] YAPAR G, LEVIE R, KUTYNIOK G, et al. Real-time outdoor localization using radio maps: A deep learning approach[J]. IEEE Transactions on Wireless Communications, 2023, 22(12): 9703-9717.
- [10] 徐亚运, 王波, 赵耀东. 城市电磁环境可视化及系统设计研究[J]. 现代电子技术, 2023, 46(10): 154-158.
- XU Y Y, WANG B, ZHAO Y D. Research on visualization of system of urban electromagnetic environment[J]. Modern Electronics Technique, 2023, 46(10): 154-158.
- [11] 刘继, 杨金瑞. 基于网格近邻优化的密度峰值聚类算法[J]. 计算机应用研究, 2024, 41(4): 1058-1063.
- LIU J, YANG J R. Density peak clustering algorithm based on grid neighbor optimization[J]. Application Research of Computers, 2024, 41(4): 1058-1063.
- [12] JIANG L F, HUANG B Y, CHEN L F, et al. Research on location planning of 5G base station based on DBSCAN clustering algorithm[C]//2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA). Piscataway, NJ, USA: IEEE, 2023: 669-675.
- [13] ZHANG X L, ZHOU S B. WOA-DBSCAN: Application of whale optimization algorithm in DBSCAN parameter adaption[J]. IEEE Access, 2023, 11: 91861-91878.