

ZHANG Hongrui, DONG Liang, HE Lesheng, HUANG Wengeng, WANG Jing, ZAINAL ABIDIN Zamri Bin, YANG Fenghui, RAZAK Dalilah Nur Fathiah Bintin Hanim, GAO Guannan. Automated Detection Method for Solar Radio Type III Bursts Using Phase Clustering and Hierarchical Thresholding (in Chinese). *Chinese Journal of Space Science*, 2026, **46**(3): 592–603. CSTR:32142.14.cjss.2025-0134. DOI:10.11728/cjss2026.03.2025-0134

基于相位编组与多级阈值的太阳射电 Ⅲ型暴自动检测算法*

张洪瑞¹ 董亮² 何乐生¹ 黄文耿³
王晶³ ZAINAL ABIDIN Zamri Bin⁴ 杨凤辉²
RAZAK Dalilah Nur Fathiah Bintin Hanim⁴ 高冠男²

1(云南大学信息学院 昆明 650500)

2(中国科学院云南天文台 昆明 650011)

3(中国科学院国家空间科学中心 北京 100190)

4(马来亚大学理学院物理系天文与天体物理研究中心 射电宇宙学研究实验室 吉隆坡 50603)

摘要 太阳射电Ⅲ型暴是太阳活动中最为常见的射电爆发类型之一, 其具有快速频率漂移等特征, 是耀斑爆发和日冕物质抛射等太阳活动事件的示踪器。现有的识别方法面临连续爆发检测能力弱、小样本学习不足等问题, 导致误检率高和召回率低。针对这些缺陷, 提出一种基于相位编组法的自适应轮廓检测算法, 该方法融合灰度梯度相位信息与多级阈值过滤, 通过形态学操作拟合太阳射电Ⅲ型暴频谱轮廓, 实现对太阳射电Ⅲ型暴频谱的自动识别, 通过计算对应的爆发起止时间、爆发起止频率、频率漂移率等有效信息, 对海量频谱观测数据进行快速批量识别与分析。基于云南天文台澄江米波太阳射电望远镜的观测数据集开展应用, 经检验, 该方法在识别太阳射电Ⅲ型暴上实现了 93.5% 的召回率。

关键词 太阳射电Ⅲ型暴, 相位编组法, 多级阈值, 形态学, 自动识别

中图分类号 P161.4

Automated Detection Method for Solar Radio Type III Bursts Using Phase Clustering and Hierarchical Thresholding

ZHANG Hongrui¹ DONG Liang² HE Lesheng¹
HUANG Wengeng³ WANG Jing³ ZAINAL ABIDIN Zamri Bin⁴

* 云南省重点研发项目 (202303AP140003), 乌鲁木齐市“红山科创英才”集聚计划科技“小组团”人才项目 (B241018006) 和国家重点研发计划项目 (2020SKA0110202, 2022YFE0140000) 共同资助

2025-07-31 收到原稿, 2026-01-13 收到修定稿

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

YANG Fenghui² RAZAK Dalilah Nur Fathiah Bintin Hanim⁴ GAO Guannan²

1(School of Information, Yunnan University, Kunming 650500)

2(Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011)

3(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

4(Radio Cosmology Research Lab, Centre for Astronomy and Astrophysics Research, Department of Physics,
Faculty of Science, Universiti Malaya, Kuala Lumpur 50603)

Abstract Solar radio type III bursts are among the most common types of radio bursts in solar activities. They are characterized by features such as rapid frequency drift and serve as a precursor tracer for solar activity events like solar flares and coronal mass ejections. Existing identification methods face problems such as weak capability in detecting continuous bursts and insufficient learning with small samples, leading to high false detection rates and low recall rates. To address these shortcomings, this paper proposes an adaptive contour detection algorithm based on the phase grouping method. This method integrates gray gradient phase information with multi-level threshold filtering, fits the spectral contour of solar radio type III bursts through morphological operations, and realizes the automatic identification of the spectrum of solar radio type III bursts. Applied to the observation dataset from the Chengjiang Meter-wave Solar Radio Telescope of Yunnan Observatory, it has been verified that this method achieves a 93.5% recall rate in identifying solar radio type III bursts.

Key words Solar radio type III bursts, Phase grouping method, Multi-level threshold, Morphology, Automatic recognition

0 引言

作为太阳系的核心天体, 太阳活动对太阳系内的其他天体产生了显著影响. 对于地球, 太阳爆发驱动的空间天气灾害严重影响了人类的生产生活, 太阳耀斑产生的高能辐射对地球电离层的干扰会影响短波通信和导航系统^[1]; 日冕物质抛射所携带的磁场会与地球磁场相互作用, 产生的地磁暴影响到电力传输系统, 可能会导致变压器损坏和大规模停电^[2].

太阳爆发活动往往伴随着高能电子的产生, 高能电子在太阳大气和行星际中可以产生活跃的电离辐射, 其亮度最高可达 10^{15} K, 这种高亮度的短时射电增量被称为射电暴^[3]. 太阳射电暴给空间物理研究提供的独特视角, 有助于描绘更细致的爆发物理过程信息.

太阳射电 III 型暴是一种最常见的射电暴, 是太阳剧烈活动中产生的非热高能电子束在日冕和行星际空间运动的最佳示踪器^[4], 在毫秒至分钟量级的时间分辨率下反映了太阳爆发能量释放和粒子加速等物理过程, 是耀斑、日冕物质抛射等太阳爆发事件的先兆^[5]. 深入研究太阳射电 III 型暴, 开展精准频谱识别, 针对其频谱参数进行提取, 及时启动空间天气事件预

警预报工作, 降低爆发事件影响, 对保护多种现代技术系统的安全稳定运行意义重大.

随着射电望远镜技术的快速发展, 全球各地涌现出大量的太阳射电观测设备, 例如 LOFAR (Low-Frequency ARray), DART (DAocheng Radio Telescope), EOVS (Expanded Owens Valley Solar Array), e-Callisto (Compound Astronomical Low-frequency Low-cost Instrument for Spectroscopy and Transportable Observatory) 等, 这些设备获得了海量的太阳射电观测数据. 面对庞大的数据量, 依靠传统人工检查对每天的观测频谱中爆发时段频谱进行甄别和分类, 不仅耗费大量的人力和时间, 还严重依赖识别人员的专业知识和经验, 同时人工识别过程中也常会出现误判和漏判的情况^[6]. 目前高效、准确的太阳射电暴自动识别技术正在研发中, 技术成熟应用后能够极大地提高数据处理效率, 减少人为误差, 更准确、及时地捕捉射电爆发事件, 为空间天气事件的预警预报提供有力的数据支持.

迄今已出现不少对太阳射电爆发的识别研究, 并取得了一系列的成果. Lobzin 等^[7]使用灰度化和固定阈值的方式来处理图像噪声, 然后基于 Radon 变换

来检测Ⅲ型暴的直线特征,成功实现了太阳射电爆发的自动识别,但是对阈值高度敏感,且难以检测弱爆发事件. Gao^[8]采用深度学习的方法,使用 DnCNN 与 FFDNet 模型进行图像降噪,使用 FPN 作为主干网络,利用 ResNet 进行特征提取,取得了较好的降噪效果,提升了多爆发检测的能力. Shen 等^[9]采用 Yen 算法对图像进行自适应阈值二值化处理来消除图像噪声,然后基于概率霍夫变换的算法来识别近似直线的射电暴结构,并根据Ⅲ型暴频率漂移率范围($1\sim 10\text{ MHz}\cdot\text{s}^{-1}$)筛选目标,同时获取到了射电暴的起止频率、起止时间、频率漂移率等信息,但存在连续爆发的检测能力较弱的问题. Bussons 等^[10]采用深度学习的方法,基于改进的 AlexNet 构建端到端检测框架 deARCE,通过动态概率阈值优化和跨台站交叉验证,显著提升太阳射电暴检测鲁棒性,并实现多源频谱数据的多爆发联合识别. Deng 等^[11]采用深度学习的方法,基于改进的 YOLOv8 n 模型构建端到端检测框架,引入全维度动态卷积 ODConv 增强特征提取能力,创建基于 ConvNeXt 的多尺度特征融合网络 MSFF-NeXt 优化特征损失,并采用 CIoU 损失函数提升回归精度,实现了 82.4% 的平均检测精度和高速实时检测多类别射电暴与定位,但各类爆发样本数量不均,导致模型难以学习其特征.

针对现有方法在多类型噪声鲁棒性、连续爆发检测能力以及小样本学习方面存在的缺陷,研究从频谱图像本身的特征入手,提出一种基于相位编组法的自适应轮廓检测算法.该方法通过融合灰度梯度相位信息与多级动态阈值,在保留Ⅲ型暴连续频率漂移特征的同时,能有效区分外界干扰与爆发信号(例如条纹与脉冲干扰).利用云南天文台澄江米波太阳射电望远镜的观测数据进行的识别实验表明,该算法在干扰滤除与爆发事件完整性检测上均取得了较好的实验结果.

1 太阳射电Ⅲ型暴的频谱特征

太阳射电Ⅲ型暴作为爆发最为频繁的射电暴类型,具有较为突出的爆发特征.一是从高频向低频快速漂移;二是爆发持续时间较短,仅有数秒至数十秒钟.如图 1 所示.

基于上述观测特征,考虑从频谱图上的图像特征入手分析,提取频谱图中像素点差异化区域,用以从

海量的频谱数据中提取潜在的太阳射电爆发信号.

2 相位编组法

2.1 算法原理

在一幅灰度图像中,像素点的亮度值范围是 $0\sim 255$,像素点的亮度代表了该点的能量强度,相邻像素点之间亮度值的差值就形成了梯度.

相位编组法是一种基于图像局部梯度信息进行特征识别的技术,核心原理是通过分析像素点之间的相位一致性来提取连贯性结构特征.其最主要特点是:图像中特征边缘处的像素梯度方向垂直于边缘走向^[13],即该像素点的梯度方向角度与图像特征的边缘延伸趋势相互垂直.基于该特点,在突发性太阳射电爆发信号中,其在频谱图上的表现为像素点亮度的突变.这为该算法应用于爆发信号搜索提供了依据.

2.2 算法流程

相位编组法提取图像特征边缘的过程如下.

步骤 1 计算水平梯度 G_x 和垂直梯度 G_y . 水平梯度和垂直梯度分别用来衡量图像在水平方向和垂直方向灰度值的变化率,本研究在 Python 编程环境下调用 OpenCV 中 Sobel 微分算子对 G_x 与 G_y 进行求解. G_x 和 G_y 的表达式为

$$G_x = P * \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad (1)$$

$$G_y = P * \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (2)$$

其中, P 表示图像中每个像素点及其周围 8 个像素点

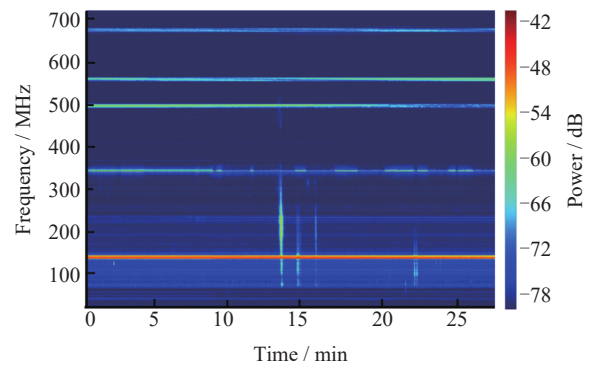


图 1 太阳射电Ⅲ型暴

Fig. 1 Solar radio type III bursts

的像素值.

步骤 2 根据 G_x 与 G_y 计算各个像素点的梯度幅值 M 和梯度方向角 θ , 提取图像上有边缘特征的像素点. M 代表各个像素点的边缘强度, 值越大, 边缘越明显; θ 代表各个像素点上梯度向量所指的方向. M 和 θ 的表达式为

$$M = \sqrt{G_x^2 + G_y^2}, \quad (3)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right). \quad (4)$$

步骤 3 编组区划分. 编组区是用于把对应角度范围内的像素点区分开来. 为了在算法实现中追求简单、高效、对称的结果, 将 360° 的梯度方向进行平均划分, 划分数量一般以 2^N 为前提. 本研究中将 N 设置为 3, 把梯度方向划分成 8 个编组区, 每个编组区对应 45° 的角度范围, 如图 2 所示.

步骤 4 将计算出的边缘像素点划分到对应的编组区. 如图 3 所示, 每个方格中的数字代表方格颜色所对应的灰度值, 利用式 (1)(2)(4) 计算得灰度值为 100 的像素点梯度方向角 $\theta = 0.3^\circ$, 对应在 $[0^\circ, 45^\circ)$ 范围内, 则可将这个像素点划分到图 2 中的 0 号编组区.

3 太阳射电Ⅲ型暴频谱识别流程

基于中国科学院云南天文台抚仙湖观测基地的米波段太阳射电观测数据, 开展太阳射电Ⅲ型暴的识别, 观测频率范围为 $70 \sim 700$ MHz. 观测设备每天进行 10 h 观测, 生成 7 张频谱图. 08:00—12:00 LT 2 张 (左右旋各一张), 12:00—16:00 LT 2 张 (左右旋各一

张), 16:00—18:00 LT 2 张 (左右旋各一张), 18:30 LT 左右, 将上述频谱图合成一张全天频谱图.

整个太阳射电Ⅲ型暴的自动识别流程如图 4 所示, 主要分为频谱图的处理和频谱图文件名的信息提取两部分, 最后根据两部分处理结果计算出分析Ⅲ型暴所需的各项参数.

3.1 频谱图预处理

以一张标准的 4 h 频谱图为例, 其图片大小为 7680×1080 pixel, 采用不同的颜色深度表示频谱强度, 如图 5 所示.

在图片处理过程中, 由于图片中局部的射电暴时刻才是需要处理的部分, 所以首先对图像进行裁剪,

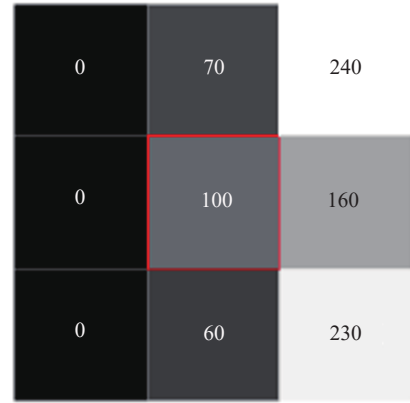


图 3 九格像素灰度图

Fig. 3 Nine-cell pixel grayscale image

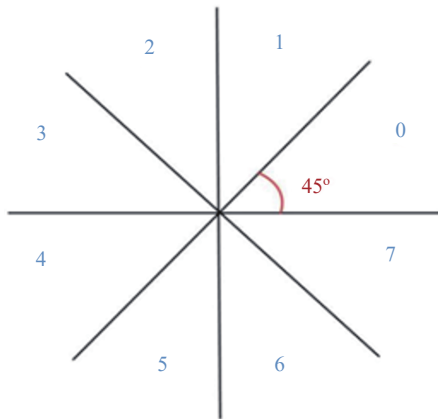


图 2 八等分编组区

Fig. 2 Eight-part grouping area

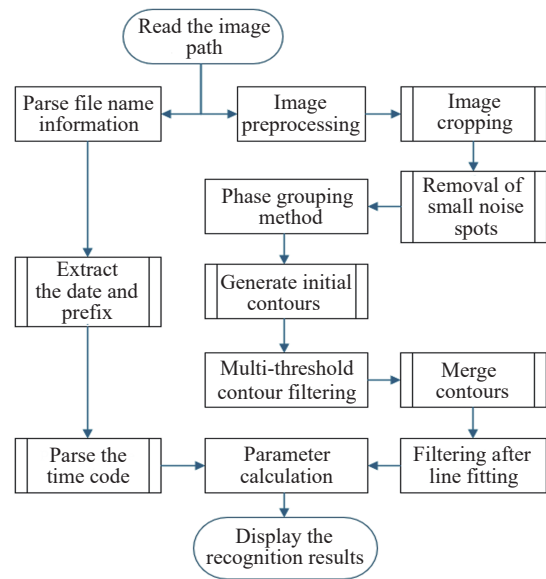


图 4 实验流程

Fig. 4 Experimental flow chart

将频谱图四周的白色边框以及其中的频率轴、时间轴和颜色强度值去除掉, 并将频谱图灰度化, 处理后的灰度图大小为 7377×975 pixel, 结果如图 6 所示。

经灰度化处理的图像由三个部分组成: 背景、干扰和射电暴。进一步基于形态学对频谱图进行降噪处理, 提高识别准确率。形态学操作是以图形形状为基础开展的一系列图像处理操作, 其中最为基础的是膨胀和腐蚀。

从数学层面而言, 膨胀和腐蚀均是将图像或图像的部分区域 (称作 A) 与结构元素 (称作 B) 进行卷积。膨胀是求局部最大值的操作, 结构元素 B 与图像 A 卷积, 计算结构元素 B 覆盖的区域像素点的最大值, 并把这个最大值赋给参考指定的像素点, 使图像中的高亮区域逐渐增长; 腐蚀则与之相反, 是求局部最小值的操作, 可减少图像中的高亮区域^[14]。腐蚀与膨胀的数学表达式为^[15]

$$A \ominus B = \{z | B_z \subseteq A\}, \quad (5)$$

$$A \oplus B = \{z | B_z \cap A \neq \emptyset\}. \quad (6)$$

其中, $\ominus$ 表示 B_z 所有元素都包含在 A 内, B_z 为把结构元素 B 平移到位置 z 后得到的区域; $\oplus$ 表示 B_z 与 A 的交集。

首先应用形态学开运算去除图像中零散的、小尺寸的高亮斑点, 这些斑点主要源于太阳射电观测中常见的瞬态噪爆或其他短时射频干扰, 该操作有助于突出 III 型暴的连续轮廓特征。开运算是先腐蚀后膨胀的过程, 其效果可以去除低亮度背景下高亮度特征的部分毛刺, 使得图像特征的边缘更为平滑。

本实验中小噪点去除部分使用 3×3 的方形结构元素进行开运算操作, 该尺寸经验证可有效抑制传感器噪声并保持信号完整性。经过形态学开运算处理后, 效果如图 7 所示, 开运算处理后的灰度图中多数零散的小干扰点被去除, 为后续识别工作排除了部分噪声干扰。

3.2 特征轮廓提取

利用 Soble 微分算子计算出每一个灰度像素点的梯度幅值和梯度方向角。

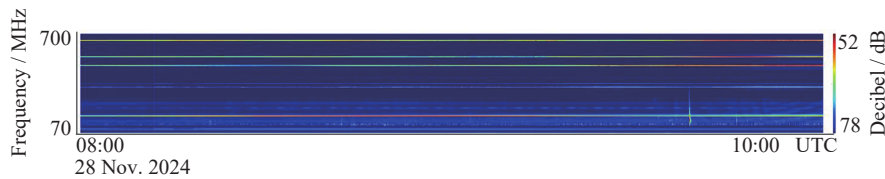


图 5 原始频谱

Fig. 5 Raw spectrum

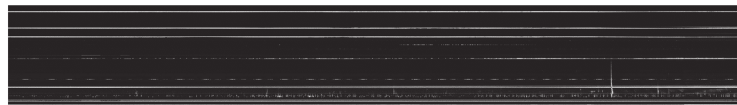


图 6 裁剪后的灰度频谱

Fig. 6 Cropped grayscale spectrum diagram

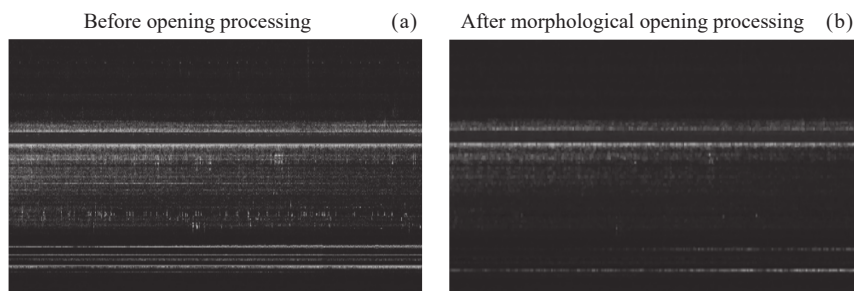


图 7 形态学开运算处理结果

Fig. 7 Morphological operation processing results

3.2.1 边缘特征像素点筛选

相位编组法最核心的特点是图像中边缘处的像素梯度方向垂直于边缘走向. 根据这一特点推断出, 一个像素点所计算出的梯度方向角越趋近于 0° 和 180° , 则其边缘特征在垂直方向上的变化趋势越明显; 所计算出的梯度方向角越趋近于 90° 和 270° , 则其边缘特征在水平方向上的变化趋势越明显. 由于太阳射电Ⅲ型暴有高频谱漂移特性, 在频谱图中趋近于垂直的图像形态, 所以可以聚焦处理 0, 3, 4, 7 编组区内的像素点, 即梯度方向角在 $[315^\circ, 45^\circ)$ 和 $[135^\circ, 225^\circ)$ 范围内的像素点.

在一张含有太阳射电Ⅲ型暴的频谱图中, Ⅲ型暴所对应灰度值是较高的部分, 为此频谱图中射电暴的边缘像素点所计算出的梯度幅值也是较大的. 本研究中保留梯度幅值大于最大梯度幅值 10% 的像素点这一动态阈值, 该设计可以 100% 去除背景部分计算出的梯度幅值, 也可以剔除部分弱边缘特征, 其判断公式为

$$M \geq M_{\max} \times 0.1. \quad (7)$$

其中, M 为梯度幅值, $M_{\max}$ 为梯度幅值的最大值.

经过编组区和梯度幅值筛选后保留的这些像素点具有与Ⅲ型暴更为相近的垂直特征, 也简化了后续操作的计算量.

3.2.2 初始特征轮廓连接

使用形态学闭运算的方法连接相位编组后保留的像素点. 闭运算是先膨胀后腐蚀的过程, 其效果能够扩充高亮度区域的范围, 有助于连接零散的拥有边缘特征的像素点, 形成初始的轮廓.

选择尺寸适宜的结构元素开展闭运算操作至关重要. 通过云南天文台澄江米波太阳射电望远镜观测数据人工筛选出 100 幅射电频谱图构建实验样本集, 其中共包含 136 个Ⅲ型暴爆发样本. 这里依据特征边缘连接正确率 R 和噪声连接错误率 E 作为连接性能评判指标, 有

$$R = \frac{N_{\text{correct}}}{N_{\text{total}}} \times 100\%, \quad (8)$$

$$E = \frac{N_{\text{error}}}{N_{\text{noise}}} \times 100\%. \quad (9)$$

其中, N_{correct} 表示正确连接的特征边缘数量, N_{total} 表示应该连接的特征边缘总数, N_{error} 表示错误连接的噪声数量, N_{noise} 表示噪声元素总数. 通过对实验样本集的统计分析, 根据上述评判指标绘制结构元素尺寸连接性能对比表, 如表 1 所示. 结果显示, 7×2 的条形核在特征边缘连接与错误连接之间达到最佳效果.

连接轮廓的过程如图 8 所示. 从图 8 可以看出, 7×2 中的两列所选取的是 7 列中第 3 和 4 列, 即结构元素的中心列和其左边一列, 使其在围组轮廓时向外扩张一个像素, 以确保射电暴轮廓的完整性. 使用

表 1 核尺寸连接性能对比
Table 1 Comparison of kernel size connection performance

核尺寸	特征边缘连接正确率/(%)	噪声连接错误率/(%)
3×2	37.5	0.9
5×2	73.5	4.6
7×2	92.6	10.3
9×2	94.8	39.8

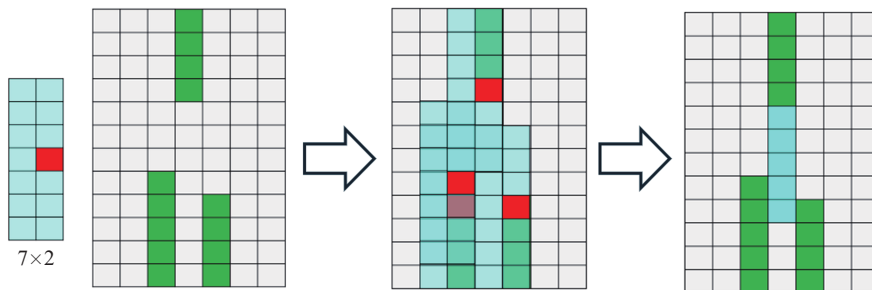


图 8 模拟形态学闭运算连接轮廓过程 (红色表示结构元素中心, 绿色为相位编组法计算后有边缘特征的各段零散轮廓)

Fig. 8 Simulated morphological closed operation connection contour process (Red represents the center of structural elements, and green represents the scattered segmented contours with edge features obtained via the phase grouping method)

7×2 的竖条形结构元素进行闭运算操作, 实现初始特征轮廓连接. 整个闭运算操作流程是用 7×2 的结构元素模拟整个连接过程, 红色的像素点表示结构元素的中心, 绿色部分为相位编组法计算后有边缘特征的各段零散轮廓, 执行膨胀操作, 用 7×2 的结构元素中心去覆盖每一个轮廓像素点, 然后执行腐蚀操作, 用 7×2 的结构元素覆盖每一个膨胀后增加的像素点, 如果结构元素能完整重合, 则保留其结构元素中心点, 删除结构元素无法完整重合的点, 最后结果将多个轮廓成功连接. 这种结构元素与Ⅲ型暴近似直线的特征结构相似, 也减少了形态学操作时对趋于水平区域的连接, 进一步降低了噪声干扰.

3.3 多级阈值轮廓过滤干扰信号

在提取特征轮廓后, 从图 9 可以看出, 除了所需

的Ⅲ型暴特征以外, 还有各种形态各异的干扰也被围组出了特征轮廓. 这些是射电望远镜在观测过程中受到的外界人为因素、自然干扰电磁波、观测仪器部件稳定性等多种因素而导致的. 在初步轮廓检测后, 研究采用基于几何特征的多级阈值筛选方法优化检测结果. 通过设计多级阈值依次过滤灰度图中具有相似特征且与Ⅲ型暴频谱特征有一定差异的各种干扰信号轮廓.

3.3.1 基于像素面积阈值的滤波

在频谱图中, 一个像素点表示一个单位面积, 每个围组出的轮廓对应的面积由其轮廓内的所有像素点组成. 通过对频谱图中射电爆轮廓面积的观察, 这里设定去除轮廓面积阈值小于 90 pixel, 用于剔除面积尺寸过小的干扰信号.

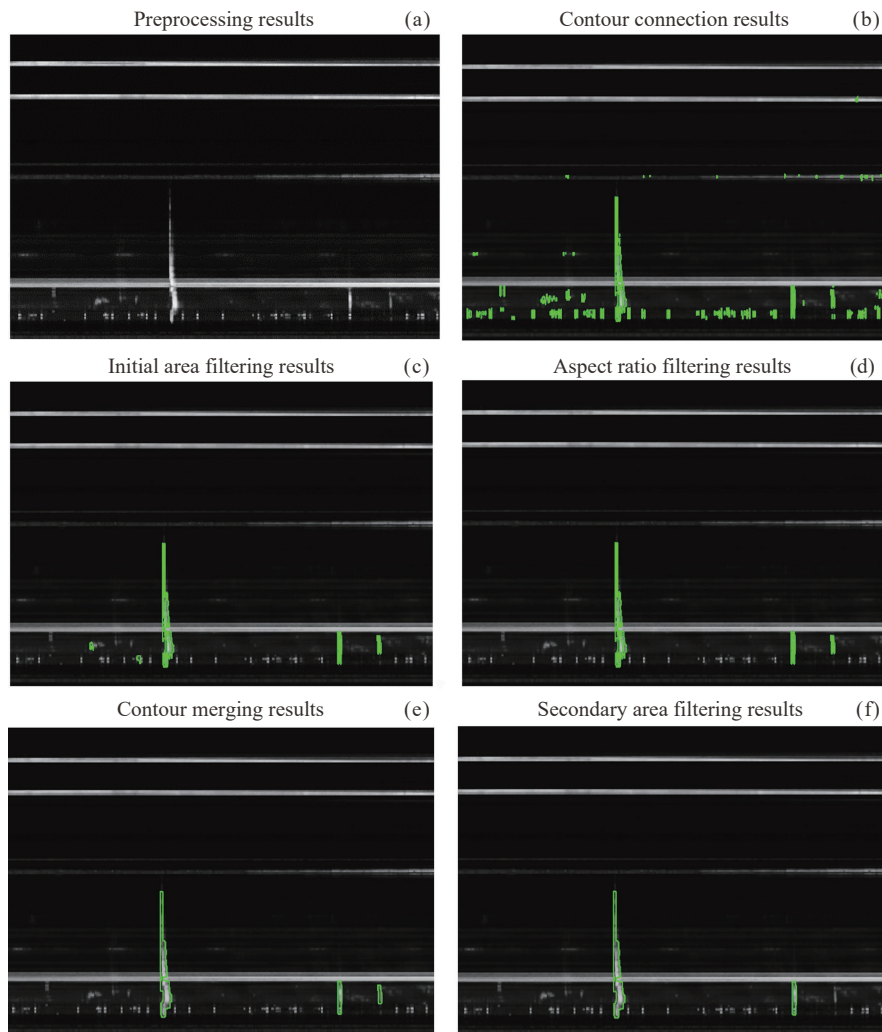


图 9 分步骤处理结果

Fig. 9 Processing results of different steps

通过对实验样本集进行统计分析, 绘制了信号频谱轮廓面积分布折线图, 如图 10 所示. 结果显示, 真实太阳射电爆发频谱的最小面积为 (97 ± 5) pixel, 而绝大部分短时干扰信号频谱产生的轮廓集中在 90 pixel 以下区域. 据此设置 90 pixel 作为面积阈值, 在保障 100% 以上信号检出率的同时有效过滤 83.1% 的噪声干扰.

3.3.2 基于宽高比形态的滤波

基于太阳射电Ⅲ型暴的物理特征分析, 其频谱呈现垂直方向扩展特性, 而干扰信号的频谱多表现为水平带状分布. 通过对实验样本集进行统计分析, 计算出图中所有太阳射电Ⅲ型暴频谱轮廓的宽高比, 其中最大的宽高比值为 0.59, 为此设定 0.6 的宽高比为区分爆发与干扰信号的判决阈值, 在确保 100% 信号轮廓能够保留的同时, 也提供了一定的余量. 其判断公式为

$$W/H \leq 0.6. \quad (10)$$

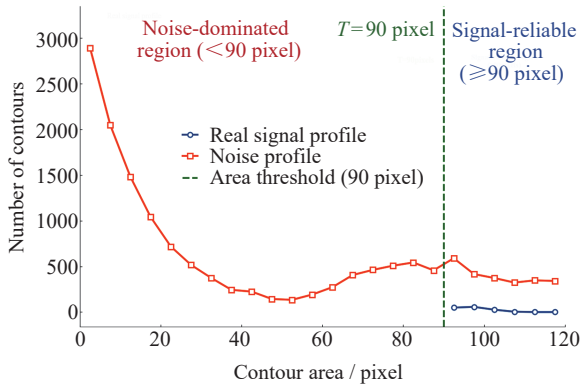


图 10 信号频谱轮廓面积分布折线

Fig. 10 Signal spectrum contour area distribution line

其中, W 为轮廓的水平宽度, H 为轮廓的垂直高度.

该宽高比过滤的处理流程如下.

(1) 在 Python 编程环境下通过调用 OpenCV 中的 `cv2.boundingRect` 函数遍历每个轮廓上的所有像素, 计算出每个轮廓对应的水平方向上最右端和最左端的像素点相减得到 W , 垂直方向上最上端和最下端的像素点相减得到 H .

(2) 通过式 (9) 判断, 筛选出符合的轮廓.

这一约束基于Ⅲ型暴信号在频谱图中的物理特性设计, 因为真实信号通常在频率轴上快速延伸而在时间轴上短时爆发.

3.3.3 基于合并轮廓面积的滤波

由于 VHF-P 波段存在大量无线电干扰信号, 其在频谱图上呈现出高亮的横条纹, 这些横条纹造成爆发结构间断, 有时甚至遮盖较弱的爆发频谱, 影响识别准确性.

通过对实验样本集进行统计分析, 人工观测灰度频谱图计算得出高亮横条纹的最大高度为 14 pixel. 这种条纹在图像中和射电暴特征的灰度值很相近, 因此可将高亮横条纹视为水平方向延伸的特征结构, 在图 8 中使用 7×2 pixel 的竖条形结构元素连接特征轮廓时, 如果一个射电暴刚好被高亮横条纹所截断, 可视为两个特征频谱交叉重合, 则连接上下轮廓的过程中连接部分的高亮横条纹特征, 也可修复断裂的射电暴特征. 7×2 pixel 的条形核能够在第一次连接轮廓特征时修复上端和下端的各 3 pixel, 如图 11(a) 所示, 该射电暴的轮廓被截断, 其剩余的断裂间隙最大还有 8 pixel, 所以本研究中设计了 9×9 pixel 的矩形核进行闭运算, 合并断裂和相邻的轮廓, 9×9 pixel 的矩形核

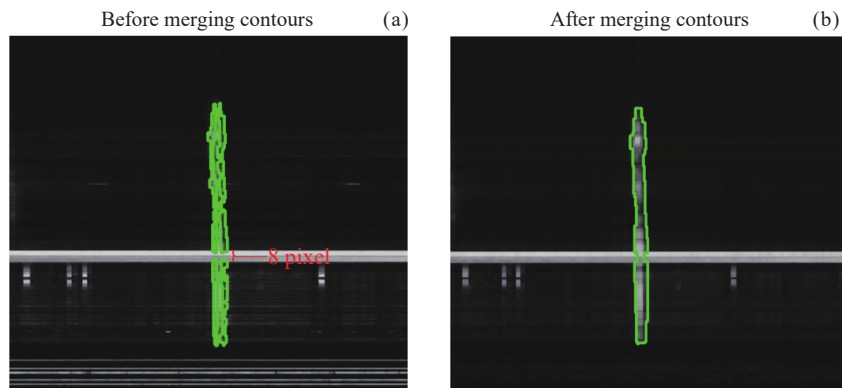


图 11 合并轮廓结果

Fig. 11 Results of merging contours

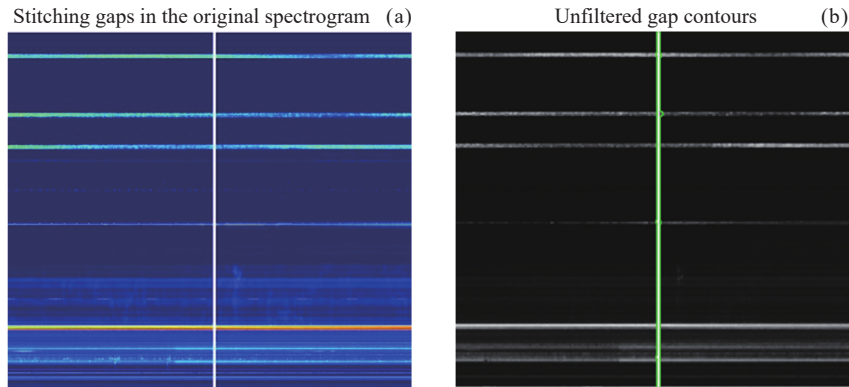


图 12 绘制频谱图过程中的拼接缝隙

Fig. 12 Splicing gaps in the spectrum drawing process

刚好能修复最大 8 pixel 的间隙,也可以合并水平方向上的相邻轮廓,形成更完整的轮廓区.实验操作结果如图 11(b) 所示,多个细小的轮廓在形态学闭运算后成功合并为一个整体,完整地围组出了射电暴的轮廓.

在第一级面积过滤时,采用 90 pixel 的面积阈值过滤后,仍有 16.9% 的面积超过 90 pixel 的噪声无法被滤除.在执行合并轮廓操作时,噪声的零散分布特点,导致大部分噪声轮廓并不会被合并.当两部分最小面积 90 pixel 的被截断Ⅲ型暴轮廓合并时,其合并后的像素面积一定大于 180 pixel,所以本研究设计了第二级面积过滤方案,滤除合并轮廓后像素面积小于 180 pixel 的轮廓,进一步滤除了像素面积在 90~180 pixel 的较大噪声.图 9 给出了多级阈值轮廓过滤的各步骤结果.

3.4 基于直线拟合长度的滤波

完成前述所有操作结束后,极少部分频谱图中的偶发性干扰仍没有被筛除,这类干扰是在绘制原始频谱图时,拼接图片时产生的缝隙,如图 12 所示.拼接缝隙的轮廓有着与Ⅲ型暴特征高度相似的图像特征,在前述操作筛选后仍会被保留下来,这里通过使用长度过滤来滤除这类干扰,基于最小二乘法拟合直线进行长度过滤,滤除频谱高度不符合太阳射电Ⅲ型暴特征的干扰图像.

3.4.1 最小二乘法直线拟合

最小二乘直线拟合是通过将轮廓点集经一系列数学计算拟合成一条直线,并获取轮廓的直线方程.在结束上述多级筛选后,剩余的轮廓在频谱图中都呈近似垂直分布,所以设拟合的直线方程为 $x = ay + b$, a 为斜率, b 为截距,其表达式为

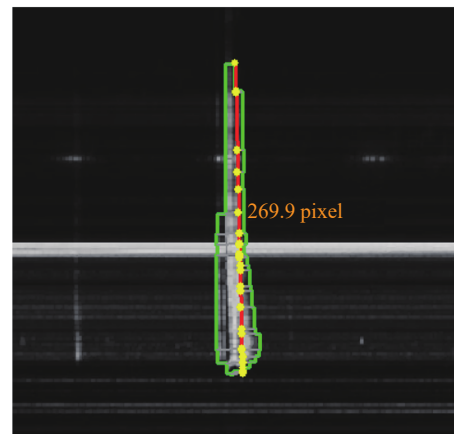


图 13 模拟拟合直线的长度计算

Fig. 13 Calculation of the length of the simulated fitting line

$$a = \frac{n \sum (x_i y_i) - \sum x_i \sum y_i}{n \sum (x_i^2) - \left(\sum x_i \right)^2}, \quad (11)$$

$$b = \frac{\sum y_i - a \sum x_i}{n}. \quad (12)$$

其中, (x_i, y_i) 为数据点集, $i=1, 2, \dots, n$.

拟合完成后,将原始点 (x_i, y_i) 正交投影到直线 $x = ay + b$ 上,计算其在拟合方向上的投影参数 t_i ,表示该点在线上位置的相对距离,表达式为

$$t_i = \frac{(x_i - b)a + y_i}{a^2 + 1}. \quad (13)$$

然后,利用投影参数的最大值与最小值之差估算轮廓在主方向上的长度 L . L 的表达式为

$$L = [\max(t_i) - \min(t_i)] \times \|v\|. \quad (14)$$

其中, $\mathbf{v} = (dx, dy) = (a, 1)$ 为拟合直线的方向向量, $\|\mathbf{v}\| = \sqrt{a^2 + 1}$ 为方向向量的模长, 用于将标量投影距离还原为图像坐标系下的真实长度. 最终结果如图 13 所示, 图中的红色直线为拟合直线, 黄色点为轮廓到直线的投影点, 拟合出的直线长度为 269.9 pixel.

3.4.2 长度过滤

通过对实验样本集进行统计分析, 直线拟合后最长的干扰轮廓长度为 975 pixel, 是绘制频谱图时的拼接缝隙, 其长度正好等于裁剪后的频谱图像素高度, 最短的Ⅲ型暴长度为 69.8 pixel. 本研究保留拟合直线长度在 65~900 pixel 长度范围内的轮廓, 来过滤较短和较长的干扰. 保留 65~900 pixel 的长度范围阈值既能最大程度地保留Ⅲ型暴轮廓, 也预留了一定的余量.

4 观测频谱图文件名解析

云南天文台澄江米波太阳射电望远镜所观测得的太阳射电频谱图的文件名中包含了对应的观测时间信息, 为此需要对观测频谱图文件名进行解析, 提取其中绘制频谱图的起始时间. 例如一幅观测频谱图的文件名为 D_253240201201_SA005006_R_20241117040000.jpg, 除扩展名以外的最后 16 位字符为计算后续参数所需信息, 其格式为 X_YYYYM-MDDHHMMSS, 对应时间是 2024 年 11 月 17 日 04:00:00 LT, 其中 X 代表的是频谱图绘制时信号的偏振方向, X 有左旋 L 和右旋 R 两种偏振.

云南天文台澄江米波太阳射电望远镜每天观测时间为 08:00—18:00 LT, 观测起始时间在 08:00 LT 和 12:00 LT 的频谱图时间长度是 4 h, 16:00 LT 的频谱图时间长度为 2 h, 转换为标准 UTC 时间, 最后 6 位数字有 000000, 040000 和 080000 三种情况. 文件名解析过程为, 删除扩展名 jpg, 提取最后 16 字符.

4.1 射电暴关键参数计算

一个射电爆发样本能够提供的有效信息包括爆发起始时间和终止时间, 爆发起始频率和终止频率, 以及频率漂移率.

具体计算过程如下: 在最终检测出的轮廓里, 以图片的左下角为原点建立直角坐标系, 每一个像素点代表一个单位长度和单位高度, 遍历所围轮廓的像素点, 找出最上端、最下端、最左端、最右端的像素点,

分别对应 $Y_{\max}, Y_{\min}, X_{\min}, X_{\max}$, 然后上下相减即为漂移频率范围, 左右相减即为爆发持续时间. 根据初始图片的坐标轴得知, 一张标准规格图片的持续时间为 4 h, 频率范围为 70~700 MHz, 根据裁剪后图片的像素值 7377×975 可计算出时间分辨率为 1.95 s, 频率分辨率为 0.65 MHz. 但是因为每天的观测时间为 10 h, 则在文件名的最后 6 位为 080000 时, 其观测时间只有 2 h. 所以在提取 16 位字符时, 如果最后 6 位为 080000, 则其时间分辨率是 0.975 s, 其余参数不变. 各参数的计算公式为

$$T_s = N_{\text{file6}} + AX_{\min}; \quad (15)$$

$$T_e = N_{\text{file6}} + AX_{\max}; \quad (16)$$

$$A = \begin{cases} 1.95 \text{ s}, N_{\text{file6}} \in [000000, 040000], \\ 0.975 \text{ s}, N_{\text{file6}} = 080000; \end{cases}$$

$$F_s = 70 + 0.65Y_{\max}; \quad (17)$$

$$F_e = 70 + 0.65Y_{\min}; \quad (18)$$

$$T_d = T_e - T_s; \quad (19)$$

$$F_d = \frac{F_s - F_e}{T_d}. \quad (20)$$

式中, N_{file6} 表示文件名最后 6 位数字, T_s 表示爆发起始时间, T_e 表示爆发终止时间, T_d 表示爆发持续时间, F_s 表示爆发起始频率, F_e 表示爆发终止频率, F_d 表示频率漂移率.

最后, 以 3.3.3 节中图片处理结果为例, 计算出的Ⅲ型暴轮廓参数从左到右依次为: 轮廓 1 的爆发起始频率为 331.95 MHz, 终止频率为 113.55 MHz, 持续时间为 29.25 s, 频率漂移率为 $-7.47 \text{ MHz} \cdot \text{s}^{-1}$; 轮廓 2 的爆发起始频率为 177.25 MHz, 终止频率为 119.40 MHz, 持续时间为 10.72 s, 频率漂移率为 $-5.39 \text{ MHz} \cdot \text{s}^{-1}$. 其中频率漂移率显示为负值说明其是从高频向低频漂移.

4.2 自动识别结果

部分识别结果如图 14 所示. 图 14(a) 中 2024 年 9 月 22 日射电暴的爆发起始频率为 290.35 MHz, 终止频率为 110.95 MHz, 持续时间为 43.88 s, 频率漂移率为 $-4.09 \text{ MHz} \cdot \text{s}^{-1}$; 图 14(b) 中 2024 年 11 月 15 日射电暴的爆发起始频率为 246.80 MHz, 终止频率为 113.55 MHz, 持续时间为 40.95 s, 频率漂

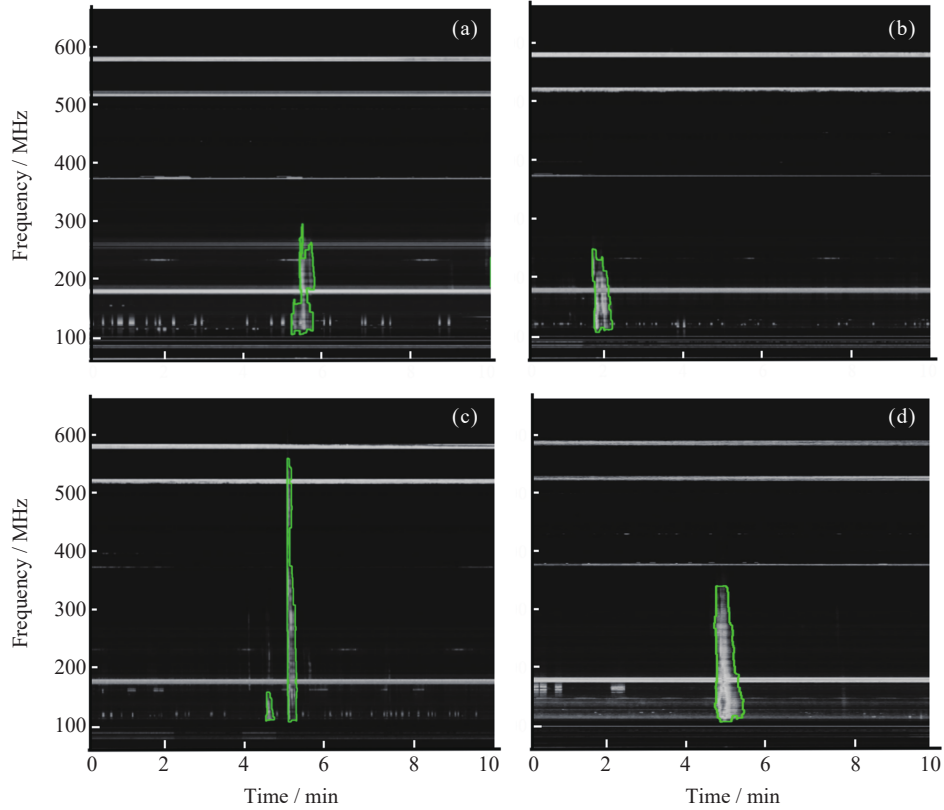


图 14 多种形态的太阳射电 III 型暴识别结果

Fig. 14 Recognition results of solar radio type III bursts with various morphologies

移率为 $-3.25 \text{ MHz}\cdot\text{s}^{-1}$; 图 14(c) 中 2024 年 11 月 18 日左侧射电暴的爆发起始频率为 162.95 MHz, 终止频率为 116.80 MHz, 持续时间为 19.50 s, 频率漂移率为 $-2.37 \text{ MHz}\cdot\text{s}^{-1}$; 右侧射电暴的爆发起始频率为 537.35 MHz, 终止频率为 115.50 MHz, 持续时间为 22.43 s, 频率漂移率为 $-18.81 \text{ MHz}\cdot\text{s}^{-1}$; 图 14(d) 中 2025 年 2 月 5 日射电暴的爆发起始频率为 330.00 MHz, 终止频率为 113.55 MHz, 持续时间为 87.75 s, 频率漂移率为 $-2.47 \text{ MHz}\cdot\text{s}^{-1}$. 以上四例识别结果证明, 该算法能够识别不同爆发位置、不同形状、不同长度的太阳射电 III 型暴, 并且能够同时识别出连续时间内的多段爆发事件.

4.3 算法性能评价

使用召回率和误检率作为识别精度评价指标, 召回率表示实际存在的特征中, 被正确识别的数量所占的比例; 误检率表示所有识别出的特征中, 被错误识别的数量所占的比例. 召回率 R_{recall} 和误检率 R_{f} 的计算公式为

$$R_{\text{recall}} = \frac{T_{\text{p}}}{T_{\text{p}} + F_{\text{n}}} \times 100\%, \quad (21)$$

$$R_{\text{f}} = \frac{F_{\text{p}}}{T_{\text{p}} + F_{\text{p}}} \times 100\%. \quad (22)$$

式中, T_{p} 表示被识别出的真实特征轮廓, F_{n} 表示未被识别出的真实特征轮廓, F_{p} 表示被识别出的错误干扰轮廓.

研究共统计了自 2024 年 9 月至 2025 年 6 月的所有射电频谱图, 总计 303 天共 1818 幅频谱图作为数据集文件. 以太阳射电 III 型暴大于其爆发中心频率 10% 的频率漂移率范围为筛选条件, 从中人工筛选出共 211 幅含有太阳射电 III 型暴的频谱图. 经实验和人工复核后, 计算得到召回率为 93.5%, 误检率为 33.8%.

5 结论

太阳射电 III 型暴作为最为常见的射电暴类型, 在空间天气研究领域有着十分重要的研究价值. 本研究设计了一种基于相位编组法的多阈值筛选太阳射电

III 型暴特征识别算法, 首先对抚仙湖观测基地所绘制的频谱图进行预处理, 经过相位编组法围组轮廓后, 通过基于几何特征的多级阈值依次滤除多种不同类型的噪声信号, 最终得到 III 型暴的轮廓结构, 并根据其轮廓以及图片文件名信息, 计算得到各项关键参数。

结果表明, 该算法可以有效地自动识别出频谱图中的太阳射电 III 型暴信号, 计算出对应的爆发起止时间、爆发起止频率、频率漂移率等有效研究信息, 并能够对海量的频谱观测数据进行快速批量识别和分析。

虽然该算法实现了较高的识别召回率, 但在降低误检率上仍有较大的进步空间, 即对频谱图进一步的噪声抑制是该研究方向下一步工作的重点。解决这一问题可以有效提升算法性能。再下一步实现软件自动处理, 在每天的观测站完成观测任务后, 将其观测结果导入软件, 可自动识别生成对应的频谱图, 有效提升了射电暴观测效率, 为空间天气研究和预警预报提供便利。

利益冲突 所有作者均声明不存在利益冲突。

参考文献

- [1] LIU Jing, WANG Wenbin, QIAN Liying, *et al.* Solar flare effects in the Earth's magnetosphere[J]. *Nature Physics*, 2021, **17**(7): 807-812
- [2] SCHRIJVER C J, KAURISTIE K, AYLWARD A D, *et al.* Understanding space weather to shield society: a global road map for 2015–2025 commissioned by COSPAR and ILWS. *Advances in Space Research*, 2015, **55**(12): 2745-2807
- [3] WILD J P. The radio-frequency line spectrum of solar noise and its interpretation in terms of fundamental physical theory. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1950, **202**(1068): 113-125
- [4] TAN Baolin, TAN Chengming, HUANG Jing, *et al.* Research progress on space very low frequency solar radio Type III bursts[J]. *Journal of Deep Space Exploration (Chinese & English)*, 2021, **8**(01): 92-99 (谭宝林, 谭程明, 黄静, 等. 空间甚低频太阳射电 III 型暴研究进展 [J]. *深空探测学报 (中英文)*, 2021, **8**(01): 92-99)
- [5] KONTAR E P, JEFFREY N L S, EMSLIE A G, *et al.* High-time-resolution observations of Type III radio bursts and their connection to solar flare particle acceleration. *Solar Physics*, 2018, **293**(10): 132
- [6] WANG Mingming, YUAN Guowu, HE Hailan, *et al.* Multi-category solar radio burst detection based on task-aligned one-stage object detection model. *Astrophysics Space Science*, 2025, **370**: 23
- [7] LOBZIN V V, CAIRNS I H, ROBINSON P A, *et al.* Automatic recognition of type III solar radio bursts: automated radio burst identification system method and first observations. *Space Weather*, 2009, **7**(4), S04002
- [8] GAO Changlin. Research on Denoising and Burst Detection Methods for Solar Radio Dynamic Spectrograms[D]. Shandong University, 2021 (高昌林. 太阳射电动态频谱图降噪与爆发检测方法研究 [D]. 山东大学, 2021)
- [9] SHEN Faxin, GAO Guannan, WANG Min. Automatic identification and parameter extraction of solar radio type III and type II bursts based on probabilistic Hough transform[J]. *Astronomical Research & Technology*, 2022, **19**(06): 559-567 (沈发新, 高冠男, 汪敏. 基于概率霍夫变换的太阳射电 III 和 II 型暴自动识别及参数提取 [J]. *天文研究与技术*, 2022, **19**(06): 559-567)
- [10] BUSSENS G J, FERNANDEZ R M, PRIETO M M, *et al.* Automatic burst detection in solar radio spectrograms using deep learning: deARCE method. *Solar Physics*, 2023, **298**: 82
- [11] DENG J, YUAN G, ZHOU H, *et al.* Real-time automated detection of multi-category solar radio bursts[J]. *Astrophysics and Space Science*, 2024, **369**(10): 99
- [12] ZHANG Juan, SHA Aimin, SUN Chaoyun, *et al.* Automatic identification of pavement cracks based on phase grouping method[J]. *China Journal of Highway and Transport*, 2008(02): 39-42 (张娟, 沙爱民, 孙朝云, 等. 基于相位编组法的路面裂缝自动识别 [J]. *中国公路学报*, 2008(02): 39-42)
- [13] KASS M, WITKIN A. Analyzing oriented patterns[J]. *Computer Vision, Graphics, and Image Processing*, 1987, **37**(3): 362-385
- [14] SONG Rui, YANG Tengyunhao, LI Yiyi, *et al.* A method for detecting defects in sanitary ceramics using dual SVM combined with morphological closing operation[J]. *Automation Application*, 2021(06): 14-17,22 (宋锐, 杨腾云浩, 李依倚, 等. 一种融合形态学闭运算的双 SVM 卫生陶瓷缺陷检测方法 [J]. *自动化应用*, 2021(06): 14-17,22)
- [15] SERRA J. Image Analysis and Mathematical Morphology [M]. London: Academic Press, 1982

作者简介



张洪瑞 男, 2000 年 7 月出生于云南省普洱市, 现为云南大学信息学院硕士研究生, 主要研究方向为太阳射电暴频谱的自动识别、无线电环境测试等。

E-mail: hongruizhang085@163.com



董亮 (通信作者) 男, 1982 年 9 月出生于四川省自贡市, 现为中国科学院云南天文台高级工程师, 硕士生导师, 主要研究方向为射电天文技术及其转化应用, 空间天气研究 (太阳射电爆发干扰导航通信预警、预报技术研究, 空间天气中生物效应的研究), 无线电环境测试技术。

E-mail: dongliang@ynao.ac.cn