

doi: 10.7690/bgzd.2018.02.013

## 基于密度聚类的改进 PRI 分选方法

王磊<sup>1</sup>, 曾维贵<sup>2</sup>

(1. 海军航空大学研究生管理大队, 山东 烟台 264001; 2. 海军航空大学科研部, 山东 烟台 264001)

**摘要:** 在雷达被动信号分选中, 对脉冲重复间隔 (pulse repetition interval, PRI) 值的提取至关重要。针对直方图统计和 PRI 变换法在脉冲重复周期值提取精度方面的不足, 提出一种基于密度聚类的 PRI 值提取方法。通过对脉冲到达时间做三级差值, 对其进行基于密度的聚类处理, 选取类内均值作为 PRI 估计值。仿真实验结果表明: 该方法提取出的 PRI 值在精确度方面明显高于传统方法, 对于存在误差抖动和杂散脉冲的雷达信号有较好的分选效果, 提高了复杂环境下被动雷达的脉冲信号分选能力。

**关键词:** 密度聚类; DBSCAN; PRI 分选; 序列搜索

**中图分类号:** TJ6   **文献标志码:** A

## Improved PRI Sorting Method Based on Density Clustering

Wang Lei<sup>1</sup>, Zeng Weigui<sup>2</sup>

(1. Brigade of Graduate, Navy Aeronautical University, Yantai 264001, China;

2. Science Research &amp; Development Department, Navy Aeronautical University, Yantai 264001, China)

**Abstract:** In the radar passive signal sorting, the extraction of the pulse repetition interval (PRI) value is critical. Aiming at the shortcomings of histogram evaluation and PRI transform method in the accuracy of pulse repetition period, a PRI value extraction method based on density clustering is proposed by this paper. By making three pulse arrival time differences, subjected to density-based clustering process, and then select the category as the mean PRI estimate. The simulation results show that the PRI value extracted by this method is significantly higher than that of the traditional method in terms of accuracy. This method has a better sorting effect on the radar signals with error jitter and spurious pulses, and improves the ability of sorting of pulsed signals of passive radar in complex environment.

**Keywords:** density clustering; DBSCAN; PRI sorting; sequence search

### 0 引言

被动雷达是根据截获目标发射的雷达信号来进行目标的探测与跟踪。相对于主动雷达系统而言, 被动雷达具有发现距离远、隐蔽性好、突防能力强等优点, 被广泛应用于军事装备中<sup>[1]</sup>。信号分选是被动雷达目标探测的关键环节, 将来自不同目标的雷达信号从相互交叠的脉冲序列中分离开。分选方法主要包括根据脉冲宽度、载波频率、脉冲幅度、到达角等特征参数进行的预分选, 以及根据脉冲到达时间进行的基于脉冲重复周期 PRI 的主分选。其中脉冲重复周期 PRI 是脉冲序列最重要的特性参数, 基于 PRI 的信号分选也是国内外学者研究最多的领域。

基于 PRI 的分选方法主要有: 序列搜索法<sup>[2]</sup>、累积差直方图法 CDIF、序列差直方图法 SDIF、PRI 变换法<sup>[3]</sup>等。序列搜索法适用于信号简单、脉冲丢

失少的情况, 对复杂环境下的信号分选能力不足; 累积差直方图法对到达时间差值做直方图统计, 适用于目标较少、脉冲丢失少的情况, 无法抑制谐波; 序列差直方图法在累积差直方图法的基础上进行了改进, 加快了运算速度; PRI 变换法加入了相位因子, 能较好地抑制谐波<sup>[4-6]</sup>。无论是 CDIF、SDIF 还是 PRI 变换法, 都是基于 PRI 区间进行的统计处理, 估值精度有限, 无法满足实际工程中对精度的需求。针对上述问题, 笔者提出一种基于密度聚类的改进 PRI 分选方法, 对脉冲到达时间累积差值进行基于密度的聚类, 能够得到较高精度的 PRI 估计值。采用边提取 PRI 值边抽取脉冲序列的方法, 对于不完全交叠脉冲序列也有较好的分选效果。

### 1 DBSCAN 密度聚类算法

密度聚类算法基本思想是, 根据一个区域点的数目来表示该区域的密度, 与某一阈值进行比较,

收稿日期: 2017-11-03; 修回日期: 2017-11-29

作者简介: 王磊(1992—), 男, 河南人, 在读硕士, 从事雷达被动信号分析与处理研究。

大于密度阈值就将其归入相近的聚类中。相比于基于距离的聚类算法，基于密度的聚类算法依据是区域密度，不限于类圆形的聚类，能够对任意形状的区域进行聚类，且对数据输入顺序及噪声不敏感<sup>[7]</sup>。

基于密度的聚类算法最典型的的就是 DBSCAN 算法，将密度相连的点的最大集合定义为簇，能够将密度聚集区从噪声中提取出来。相比于 K-means 聚类<sup>[8]</sup>，DBSCAN 算法不需要事先确定分类数，且可以发现任意形状的一类。

该方法有 3 个输入参数，待聚类数据集  $D$ 、邻域半径  $\varepsilon$ 、密度阈值  $\text{minPts}$ 。设  $x \in D$

$$N_\varepsilon(x) = \{y \in D : \text{dist}(y, x) < \varepsilon\}. \quad (1)$$

$N_\varepsilon(x)$  为  $x$  的  $\varepsilon$  邻域，定义密度

$$\rho(x) = |N_\varepsilon(x)|. \quad (2)$$

若点  $x_i$  的密度  $\rho(x_i) > \text{minPts}$ ，则称该点为核心对象，否则称为非核心对象或噪声。设定  $D_c$  为所有核心对象的集合，若  $x \in D_c$ ， $y \in N_\varepsilon(x)$ ，则称  $y$  是  $x$  直接密度可达的；设  $P_{(1)}, P_{(2)}, \dots, P_{(m)} \in D$ ，其中  $m > 2$ ，若满足  $P_{(i+1)}$  从  $P_{(i)}$  直接密度可达， $i = 1, 2, \dots, m-1$ ，则称  $P_m$  是从  $P_{(1)}$  密度可达；设  $x, y, z \in D$ ， $y$  和  $z$  分别和  $x$  密度可达，则称  $y$  和  $z$  密度相连。

从数据集  $D$  中取一未被标记的点，如果该点是核心对象，则找出所有与该点密度可达的对象，形成一个簇；否则将该点标记为噪声后，继续对其他未标记的点进行上述操作，直到遍历集合  $D$  内所有点<sup>[9]</sup>。

## 2 基于密度聚类的改进 PRI 分选方法

### 2.1 基于密度聚类的 PRI 值提取方法

设定脉冲到达时间序列为  $\text{TOA}[n]$ ， $\text{TOA}[1]$  表示第 1 个脉冲的到达时间， $\text{TOA}[n]$  表示第  $n$  个脉冲的到达时间。对脉冲序列到达时间做三级差值：

$$\text{Arr}_1 = \text{TOA}[2] - \text{TOA}[1], \dots, \text{TOA}[n] - \text{TOA}[n-1]; \quad (3)$$

$$\text{Arr}_2 = \text{TOA}[3] - \text{TOA}[1], \dots, \text{TOA}[n] - \text{TOA}[n-2]; \quad (4)$$

$$\text{Arr}_3 = \text{TOA}[4] - \text{TOA}[1], \dots, \text{TOA}[n] - \text{TOA}[n-3]; \quad (5)$$

$$\text{Arr} = \text{Arr}_1 + \text{Arr}_2 + \text{Arr}_3. \quad (6)$$

以 3 组脉冲叠加为例，PRI 值分别设定为 53.6、175.4 和 543  $\mu\text{s}$ ，噪声水平 2  $\mu\text{s}$ 。将交叠序列到达时间做三级差值后，差值分布如图 1 所示。

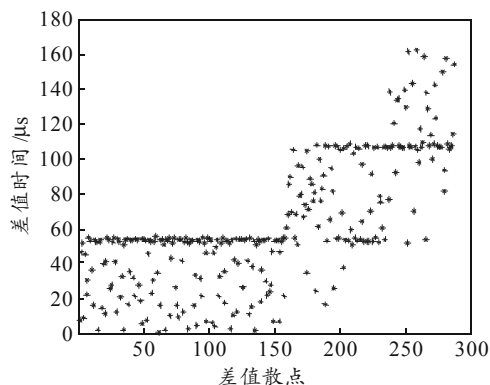


图 1 TOA 三级差值分布

由图中可以看出：在 55 和 110  $\mu\text{s}$  附近差值序列呈现明显的聚集状态，因此，可以对差值序列做基于密度的聚类处理。

应用 DBSCAN 算法，选择邻域  $\varepsilon = 1 \mu\text{s}$ ， $D = \text{Arr}$ 。假定脉冲数据流中没有周期信号出现，差值序列将呈现均匀分布，此种情况下在 2 倍  $\varepsilon$  的区间内出现的差值个数应为  $2 \cdot \text{Length}(\text{Arr}) / \max(\text{Arr})$ 。经过实际验证结果可知选择  $\text{minPts} = 6 \cdot \text{Length}(\text{Arr}) / \max(\text{Arr})$  时聚类效果最好<sup>[10]</sup>。聚类后剔除噪声点，剩余 2 个类如图 2 所示。

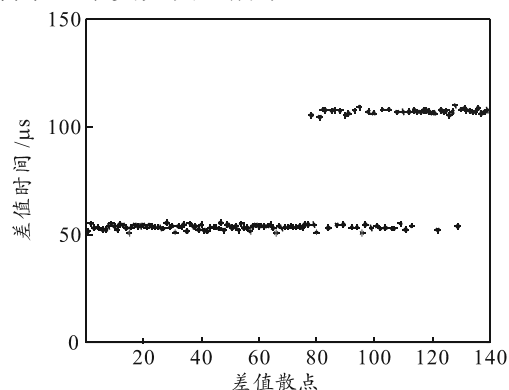


图 2 密度聚类后的差值分布

笔者从所有类中选取类内数目最大的类，将其类内均值作为暂定的 PRI 值。由于存在谐波干扰，有可能类内数目最大的类其中心不是真实 PRI 值，而是 2 倍或 3 倍谐波；因此，笔者取出暂定的 PRI 值后配合进行谐波校验，主要是检验其 1/2 和 1/3 附近是否存在类内数目相当的聚集区。如果有且满足一定条件，则说明该暂定 PRI 值为 2 次或 3 次谐波，并相应调整 PRI 值。

### 2.2 序列搜索法提取脉冲串

提取出 PRI 值之后，最重要的就是把符合该 PRI 的脉冲串从交叠脉冲序列中分离出来，才能进行下一步的信号处理分析，原始交叠的脉冲序列也

可以得到简化，方便之后的 PRI 估计和脉冲提取。

传统的序列搜索法采用试探的方式，首先选取第 2 个到达时间与第 1 个到达时间的差值作为 PRI 值，然后往后判断在 2 倍 PRI、3 倍 PRI 处是否均有脉冲，是，则继续往下直到全部找出；否，则以第 3 个到达时间与第 1 个到达时间的差值作为 PRI 值，继续上述搜索，以此类推。该方法只适用于目标较少的情况，抗抖动性差。

这里应用 2.2 节中提取的 PRI 值来进行序列搜索，从而提取相应的目标脉冲序列。判定准则为：连续找到 4 个符合 PRI 误差范围的脉冲即认为找到脉冲串开头；超过 3 倍 PRI 值仍找不到符合条件的脉冲则认为该脉冲串结束。具体步骤如下：

- 1) 以第 1 个脉冲到达时间为起点。
- 2) 判断在 1 倍 PRI、2 倍 PRI、3 倍 PRI 处是否都有满足误差允许范围的脉冲，有则认为脉冲串起始部分已找到，转入 4)，否则进行 3)。
- 3) 起点向后移一个脉冲，转入 2)。
- 4) 从起点开始根据估计出的 PRI 寻找下一个脉冲，若找到，将起点变为该脉冲，继续寻找下一个满足 PRI 的脉冲，否则转入 5)。
- 5) 若在 1 倍 PRI 处找不到脉冲，则继续向后找 2 倍、3 倍 PRI 处是否有满足误差范围的脉冲，若有继续 4)，否则认为脉冲串结束，终止本次搜索，转入 6)。
- 6) 将成功搜索到的脉冲序列从原始脉冲流中提取出来。

这种方法相较于传统的序列搜索法有更高的搜索效率和成功率。

### 2.3 算法总体实现

笔者对原有基于 PRI 的分选方法改进，提出边 PRI 值提取边进行序列搜索的方法。首先对原始脉冲数据流进行三级 TOA 差值，用 DBSCAN 算法对差值序列进行基于密度的聚类，将类内数据最多的类中心作为 PRI 估计值，配合谐波校验。被提取出的 PRI 值肯定是周期较短、出现频率较高的脉冲序列，进而依据 PRI 估计值进行序列，将个数最多的脉冲序列从原始交叠脉冲流中提取出来，原始脉冲流得到稀释。再对剩余的脉冲数据流重复进行上述过程，一次只提取一组脉冲序列，直到没有足够的脉冲可供分选。算法总体流程如图 3 所示。

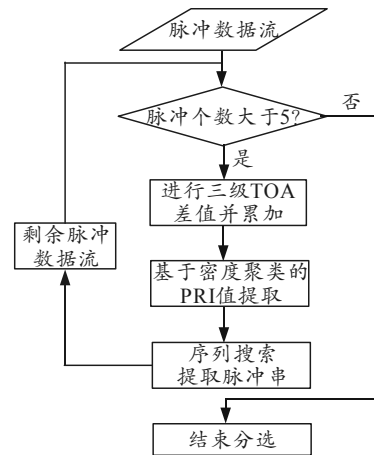


图 3 算法实现流程

### 3 仿真分析

为验证本方法的有效性，用 Matlab 进行仿真实验，设定 4 组相互交叠的脉冲序列，重复周期  $PRI1 = 53.6 \mu s$ ， $PRI2 = 175.4 \mu s$ ， $PRI3 = 543 \mu s$ ， $PRI4 = 666.7 \mu s$ ，仿真时间设定为  $6000 \mu s$ ，加入脉冲到达时间误差  $2 \mu s$ 。分别用基于密度聚类的 PRI 分选、CDIF 法和 PRI 变换法对交叠脉冲进行 PRI 值提取，并计算应用不同方法得到的 PRI 估计值误差率。

表 1 不同分选方法得到的 PRI 估计值及误差率

分选方法	第 1 组	第 2 组	第 3 组	第 4 组	误差率/%
	PRI1/ $\mu s$	PRI2/ $\mu s$	PRI3/ $\mu s$	PRI4/ $\mu s$	
CDIF 法	53.00	174.00	542.00	666.00	0.552
PRI 变换	52.24	176.60	544.80	664.20	0.982
文中算法	53.61	175.28	542.75	666.74	0.035

从表 1 中可以看出：用基于密度聚类的 PRI 分选方法提取出的 PRI 值，其精确度远高于 CDIF、PRI 变换法。基于密度聚类的 PRI 分选方法相对于 CDIF、PRI 变换法，其优势在于不用划分区间对差值序列做直方图统计，而是采用密度聚类的方法，提取出密度较大、数目最多的差值集聚区，对其做类内均值。该方法不受 CDIF、PRI 变换法统计区间取值大小的影响，做均值后能够大大减小所受噪声的影响，从而得到较高精度的 PRI 估计值。

由于实际环境中会受到各种因素干扰，各种杂散信号会进入雷达接收机。且由于雷达辐射源波束在不断扫描，所以脉冲数据流一般不会出现完全交叠的情况，往往存在部分交叠，并且在脉冲数据流中加入脉冲总个数 20% 的杂散噪声。为模拟现实不

完全交叠情况,用不同幅度区别显示各组,设定第1组脉冲分布在1000~4000  $\mu\text{s}$ 处,幅度0.5 dB;第2组脉冲分布在2000~5000  $\mu\text{s}$ 处,幅度0.75 dB;第3组脉冲分布在0~6000  $\mu\text{s}$ 处,幅度1 dB;第4组脉冲分布在0~6000  $\mu\text{s}$ 处,幅度1.25 dB,加入脉冲到时间误差2  $\mu\text{s}$ 。图4为混合交叠的待分选脉冲流,图5为仿真分选结果。

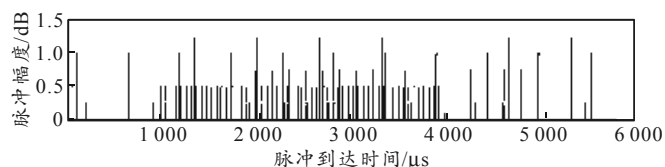
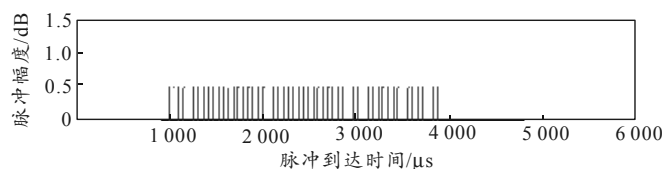
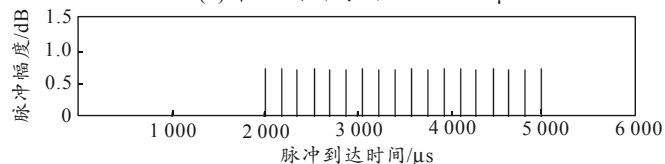


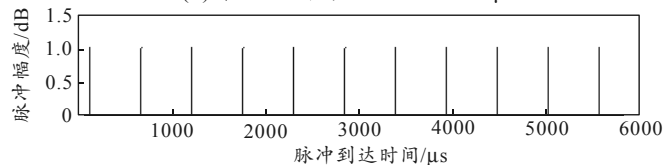
图4 混合交叠的待分选脉冲流



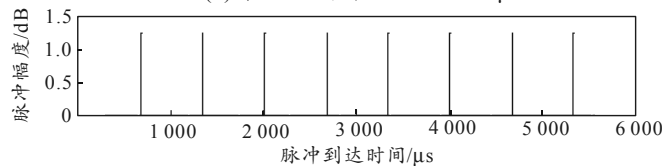
(a) 第1组脉冲序列PRI1=53.6  $\mu\text{s}$



(b) 第2组脉冲序列PRI2=175.4  $\mu\text{s}$



(c) 第3组脉冲序列PRI3=543  $\mu\text{s}$



(d) 第4组脉冲序列PRI4=666.7  $\mu\text{s}$

图5 仿真分选结果

仿真实验结果表明:笔者提出的基于密度聚类

PRI分选方法对于PRI值的提取有着较高的精确度,能够实现对交叠脉冲流的准确分选。

#### 4 结束语

笔者提出的基于密度聚类的PRI值提取方法不受类似于CDIF法的直方图统计区间、PRI变换法的盒宽度的影响,PRI提取值有着较高的精确度。笔者采用边提取PRI值边抽取脉冲串的思想,先将出现最多的脉冲提取出来,从而使原始脉冲流得到明显稀释,为后续的分选降低了难度,提高了整体分选成功率。同时也注意到,基于密度聚类的PRI值提取方法,对于DBSCAN算法的参数密度阈值minPts和邻域 $\epsilon$ 较为敏感,需要进一步提升算法性能。

#### 参考文献:

- [1] 王海峰,吴宏宇.被动雷达导引头发展历程及技术综述[J].飞航导弹,2013,43(1):78-80.
- [2] 王宇.未知雷达信号PRI的快速分选识别算法研究[D].西安:西安电子科技大学,2010.
- [3] 邹顺,张群飞,靳学明.基于PRI变换的雷达信号分选[J].计算机仿真,2006,23(6):41-44.
- [4] 李英达,肖立志,于洋,等.一种新的PRI变换分选算法[J].科学技术与工程,2014,14(12):209-213.
- [5] 邓方艺,王鹏飞.被动雷达导引头信号分选算法研究及仿真[J].现代电子技术,2012,35(13):72-74.
- [6] 关一夫,张国毅,刘志鹏.一种基于脉冲样本序列的PRI周期信号分选算法[J].电讯技术,2014,57(7):915-920.
- [7] 张晓,张媛媛,高阳,等.一种基于密度的快速聚类方法[J].数据采集与处理,2015,30(4):888-895.
- [8] 王千,王成,冯振元,等.K-means聚类算法研究综述[J].电子设计工程,2012,20(7):21-24.
- [9] KUMAR K M, REDDY A R M. A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method[J]. Pattern Recognition, 2016, 58: 39-48.
- [10] 周红芳,王鹏.DBSCAN算法中参数自适应确定方法的研究[J].西安理工大学学报,2012,28(3):289-292.