

一种改进的基于 SDIF 的信号分选算法研究

夏亦佳¹ 崔炳喆¹ 刘广君¹
XIA Yijia CUI Bingzhe LIU Guangjun

摘要

针对信号分选计算量大、实时性不强的问题,文章提出一种动态 K-means 与 SDIF 相结合的分选方法。该算法首先通过预分选将接收到的脉冲描述字(PDW)进行预分类,以此来降低脉冲序列的信号密度,并使用动态 K-means 算法进行该步骤处理。预分选处理后有效降低了主分选数据交错的复杂度、提升主分选的运算效率;主分选算法使用经典的序列差直方图法(SDIF),通过计算相邻脉冲的 TOA 差构建第一级差值直方图,再通过序列搜索,分选出所有可能的辐射源脉冲序列。仿真结果表明,该算法具有较高的分选精度和鲁棒性以及较低的运算量,适合工程化实现。

关键词

动态聚类; SDIF; 信号分选; 序列搜索; 脉冲重复周期

doi: 10.3969/j.issn.1672-9528.2025.05.028

0 引言

信号分选是电磁环境侦查的重要模块,在早期的电磁环境中,辐射源个数不多、脉冲流密度低,主要采用的是基于到达时间(TOA)的脉冲重复周期(PRI)分选。但随着技术的不断更新,电磁环境更为复杂、辐射源个数急剧增加、脉冲个数可达每秒几百万以上,因此 PRI 分选已无法满足实际需求。当前的辐射源信号分选主要采用多参数预分选和 PRI 主分选相结合的模式。其中,多参数预分选根据每个脉冲多维参数的相似性和差异性进行脉冲归类,再对每一类中的脉冲序列进行 PRI 分选,从而大大提高 PRI 的分选效率。先对脉冲描述字(PDW)中的参数进行聚类分选,可以对复杂电磁环境中的脉冲进行稀释,以便后续 PRI 分选。

在当前众多学者研究的众多信号分选算法中, PRI 分选是其中重要的一个环节,基于 PRI 分选也是各种分选方法中研究最多的方法之一。Mardia^[1]通过研究 PRI 规律,发现通过 TOA 的级差分布统计,可以分选出脉冲序列,提出了累积差分直方图法(CDIF),通过检测 PRI 和 2 倍的 PRI 进行序列搜索,可以抑制谐波,但缺点在于运算量太大。Milojevic 等人^[2]通过对 CDIF 进行改进,研究了序列差分直方图法(SDIF)。SDIF 法通过计算直方图中的当前差,对统计结果不进行累积,因此降低了计算量,提高了运算时间。此后,学者们对 PRI 分选算法提出了一系列改进,刘建新^[3]提出了采用 SDIF 法和 PRI 变换法的联合分选方法;姜勤波等人^[4]采用结合 FFT 算法和滤波技术以及过零检测提取周期序列的方正弦波插值法;向娴等人^[5]提出了通过改进的距离法提出

信号序列中的噪声后,利用网格映射,根据网格密度实现对未知辐射源信号的分选的方法。

本文提出一种基于 SDIF 的信号分选算法,采用多参数预分选法结合 SDIF 主分选算法,通过利用载频、脉宽等参数进行聚类,稀释接收到的密集脉冲流,再通过改进的 SDIF 法进行 PRI 分选,从而得到输入脉冲信号参数,降低计算量的同时,具有较高的实时性和准确率,适合工程化应用。

1 信号的特征参数

1.1 信号的脉冲描述字

常见的信号 PDW 参数有脉冲信号幅度(PA)、载频频率(CF)、到达时间(TOA)、脉冲宽度(PW)、到达角(AOA)以及脉冲重复周期(PRI)^[6]。

PA 表示接收到的脉冲功率,但易受到外界环境等因素影响,如:脉冲幅度调制、天线扫描规律等,容易带来较大的误差,因此通常不作为分选的特征参数,但可作为脉冲参数是否可信的判断指标。

PW 表示脉冲结束时刻与到达时刻之间的时间差。

CF 表示信号分选的重要参数,随着现代雷达的抗干扰能力和目标检测能力不断提升,载频的形式愈加多样。常见的有频率捷变及频率分集等。频率参数主要取决于辐射源载频的变化能力。PW 和 CF 一般结合使用,用于参数空间的分割或特征提取。

脉冲到达时间 TOA 表征的是脉冲前沿到达的时刻,是最直观描述脉冲序列的参数,根据脉冲序列的 TOA,可以计算出脉冲的 PRI 参数^[7],从而更好地区分不同的脉冲序列。

AOA 取决于辐射源和接收机之间的相对角度,可以作为

1. 中国空空导弹研究院 河南洛阳 471009

信号分选的参数,也可以协助完成辐射源测向工作。

1.2 脉间调制类型

常见的脉间调制类型有:固定重复周期、抖动重复周期、参差重复周期、滑变重复周期^[8]。

固定 PRI 的信号任意脉冲间的重复间隔是相同的。

抖动 PRI 信号的重复频率在围绕中心值的一定范围内随机变化,一般为 $\pm 10\%$ 以内。表达式为:

$$PRI_n(t) = PRI_0 + \beta \times \text{rand} \quad (1)$$

式中: β 表示抖动力; PRI_0 表示基础 PRI。

参差 PRI 信号包含一组重复周期,在脉冲间循环出现,因此该信号序列是由顺序、重复地利用 PRI 集合中的 PRI 值产生的。参差重频脉冲信号的数学表达式为:

$$PRI_n(t) = PRI_0 + \tau \times \text{mod}(k + (i - 1) \times g) \quad (2)$$

式中: τ 为脉宽; k 、 g 均为正整数; PRI_0 为基础 PRI。

滑变 PRI 是指信号的 PRI 按某种规律连续的变化,如三角波、锯齿波等。

2 基于 SDIF 的信号分选算法

2.1 总体设计方案

信号分选是从交错的脉冲序列中提取出不同辐射源对应的不同信号集合,从而获得每个辐射源对应的参数信息,主要由预分选和主分选两部分组成,首先对信号进行高精度参数检测,获得 PDW 信息;利用多参数进行预分选,对密集脉冲流进行稀释,再进行基于 PRI 的主分选,总体流程如图 1 所示。

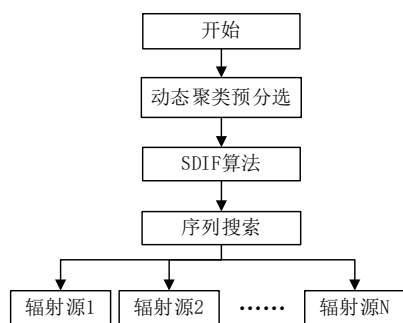


图 1 分选方案设计

本文采用动态 K-means 聚类算法对信号脉冲列进行初步分选,然后对预分选输出的脉冲序列进行 SDIF 主分选计算,得到目标雷达脉冲序列。

2.2 动态 K-means 算法

动态 K-means 是一种无监督的学习方法^[9],通过不断迭代,将数据集中在某些方面相似的数据成员进行分类组织。首先随机选取数据集中的元素作为聚类中心,计算其余元素与该中心的距离,若距离小于门限则认为是同一类,否

则作为第二个聚类中心,继续分配其余的元素。每分配一个元素,会根据该类中现有的元素重新计算聚类中心。直至所有元素被分配完毕,并且误差平方和最小则判定聚类结束。

假设检测到的 PDW 数组为 θ ,其表达式为:

$$\theta = [P_1, P_2, \dots, P_N] \quad (3)$$

式中: P 是每个脉冲描述字,表达式为:

$$P_j = [TOA_j, PW_j, AOA_j, CF_j] \quad (4)$$

聚类算法处理时,以 PW、AOA、CF 为参考标准,分别对检测到的 PDW 数据进行聚类。算法步骤如下:

步骤 1: 输入脉冲数目为 n ,脉冲数据记为 $q_1, q_2, \dots, q_n$ 。

步骤 2: 选取第一个脉冲数据 q_1 作为初始聚类中心 $P(1)$,聚类距离门限为 d 。

步骤 3: 计算第二个数据 q_2 与 $P(1)$ 之间的距离 d_{21} ,若大于距离门限 d ,则将 q_2 作为第二个聚类中心,记为 $P(2)$,反之,将 q_2 分配给类 $P(1)$,并且重新计算该类的聚类中心为:

$$P(1) = \frac{P(1) + q_2}{2} \quad (5)$$

步骤 4: 考虑第 j 个数据 q_j ,对于 $P(1), P(2), \dots, P(M)$,共 M 个聚类中心,计算 q_j 与每个聚类中心 $P(i)$ 的距离,若满足 $P(i) - d < q_j < P(i) + d$,则该数据属于第 i 个类别,否则,该数据则记为第 $M+1$ 个类别,即 $P(M+1)$ 。

步骤 5: 将所有脉冲数据分类完毕并且误差平方和最小,则认为聚类结束。

2.3 SDIF 算法

序列差值直方图法(SDIF)是在累积差值直方图算法(CDIF)的基础上进一步优化的算法^[10],相比于 CDIF、SDIF 法,可以减小计算量并且有更好的谐波抑制效果。

SDIF 算法流程图如图 2 所示。

其具体步骤如下:

步骤 1: 首先将脉冲进行到达时间 TOA 排序,计算相邻脉冲时间间隔,得到一阶差值直方图。

步骤 2: 进行谐波校验和直方图峰值是否超过门限进行判别。

步骤 3: 对直方图峰值高于检测门限 PRI 进行序列搜索。若有不止一个 PRI 对应的直方图峰值高于门限,说明脉冲序列中存在潜在的其他辐射源信号,则不能开展序列搜索,而是进行二阶差值直方图的计算,重复进行门限判别。

步骤 4: 对步骤 3 序列搜索后剩余的脉冲列计算新的一阶差值直方图,重复以上步骤,直到剩余脉冲不足以形成新的脉冲序列为止。

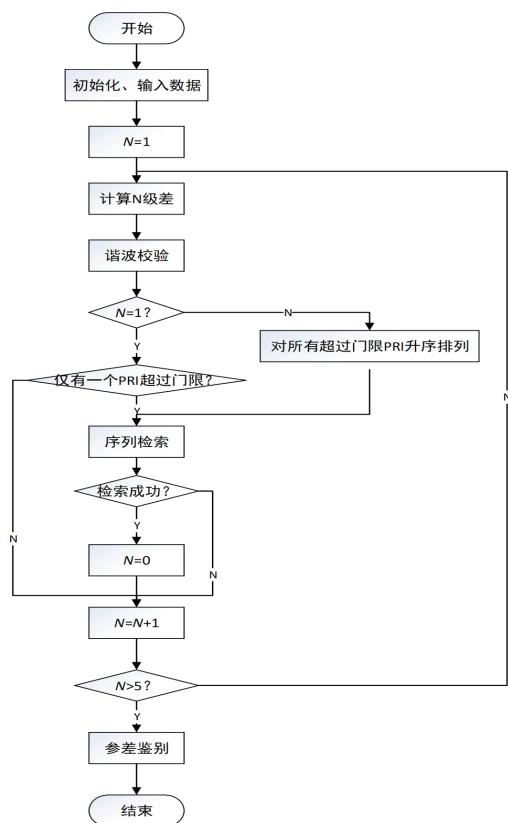


图2 序列差值直方图 (SDIF) 算法流程图

检测门限是 SDIF 算法中关键的设置参量^[11]。已知当信号脉冲的间隔时间越长,则在一定时间内,接收到的脉冲个数就会越少,因此对该信号统计的直方图值和 PRI 成反比,门限的设置应当随脉冲数增大而提高,随相邻脉冲到达时间间隔降低,其计算方法为:

$$P(\tau) = \frac{xN}{\tau} \quad (6)$$

式中: N 表示脉冲总数; τ 表示脉冲时间间隔; x 表示不大于 1 的系数。

当多个辐射源信号存在时,接收端采集到的信号脉冲间隔服从 Poisson 分布,在时间间隔 $\tau = t_2 - t_1$ 内,有 i 个随机 Poisson 出现概率符合:

$$P_i(\tau) = \frac{(\lambda x)^i}{i!} e^{-\lambda x}, i = 0, 1, \dots \quad (7)$$

式中: $\lambda = \frac{m}{T}$; T 表示采样时长; m 表示存在的脉冲子间隔数。

SDIF 的一阶差计算方法为: 相邻脉冲到达时间相减,因此一阶差直方图可以表示为:

$$P_0(\tau) = e^{-\lambda \tau} \quad (8)$$

随机事件的概率分布可近似用直方图表示,因此 SDIF 算法的高阶差直方图也是随指数分布的。假设第 D 阶差值的直方图的脉冲数为 $(N-D)^{[12]}$, 所以门限函数可表示为:

$$T_{\text{threshold}}(\tau) = x(N-D)e^{-\tau/kT} \quad (9)$$

2.4 序列搜索

序列搜索依据计算的 PRI, 随脉冲序列进行关联的过程。可以认为由同一部常规雷达辐射源发射出的每个脉冲,服从均值为 PRI_0 、方差为 σ^2 的正态分布,即 $\text{PRI} \sim (\text{PRI}_0, \sigma^2)$, 当接收到的脉冲序列中按照 PRI 搜索后,关联到 n 个脉冲后,则判定成功检测到该辐射源。假设这 n 个脉冲信号为 $\xi_1, \xi_2, \dots, \xi_n$, 由于 $\xi_i \sim (\text{PRI}_0, \sigma^2)$, 所以:

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i \sim \left(\text{PRI}_0, \left(\frac{\sigma}{\sqrt{n}} \right)^2 \right) \quad (10)$$

因变量方差 σ^2 未知,需完成均值检验即 PRI 是否等于 PRI_0 。假设 $H_0: \text{PRI} = \text{PRI}_0$, 以及该假设的对立面 $H_1: \text{PRI} \neq \text{PRI}_0$, 然后用统计方法判断哪个假设成立^[13]。选择检验统计量:

$$T = \sqrt{n-1} \frac{\bar{\xi} - \text{PRI}_0}{S} \quad (11)$$

式中: 均值 $\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$; 方差 $S^2 = \frac{1}{n} \sum_{i=1}^n (\xi_i - \bar{\xi})^2$ 。

假设 H_0 成立的条件下,统计量 T 的分布为:

$$T = \sqrt{n-1} \frac{\bar{\xi} - \text{PRI}_0}{S} \sim t_{(n-1)} \quad (12)$$

式中: $t_{(n-1)}$ 为自由度为 $n-1$ 的 t 分布。选择临界值 $t_{(n-1)}(a)$, 使得:

$$P \left\{ \left| \sqrt{n-1} \frac{\bar{\xi} - \text{PRI}_0}{S} \right| \geq t_{(n-1)}(a) \right\} = a \quad (13)$$

信号分选中,给定 a 判断 $H_0: \text{PRI} = \text{PRI}_0$ 的序列是否存在: 由具体样本 $\xi_1, \xi_2, \dots, \xi_n$ 算出 $\hat{T}$, 若 $\hat{T} > t_{(n-1)}(a)$, 则否定 H_0 , 反之则否定 H_0 。可见这一判断的置信度和 PRI 个数有关。假设在 $1-a$ 的区间估计:

$$\left(\bar{\xi} - \frac{S}{\sqrt{n-1}} t_{n-1}(a), \bar{\xi} + \frac{S}{\sqrt{n-1}} t_{n-1}(a) \right) \quad (14)$$

在容差 $\Delta = \frac{S}{\sqrt{n-1}} t_{n-1}(a)$ 时, $n = 1 + \frac{S^2}{\Delta^2} t_{(n-1)}^2(a)$ 。

3 仿真结果

3.1 动态 K-means 仿真

通过仿真模拟产生表 1 中 4 种脉冲信号。首先对脉冲信号进行聚类处理,即预分选流程。经过动态 K-means 处理的信号分布如图 3 所示。通过预分选的聚类,可以实现对信号数据的粗分类。

表 1 4 辐射源信号混合输入参数设置

辐射源序号	AOA/(°)	CF/MHz	PW/ μ s
1	15	100	10
2	10	30	14
3	8	60	4
4	47	140	20

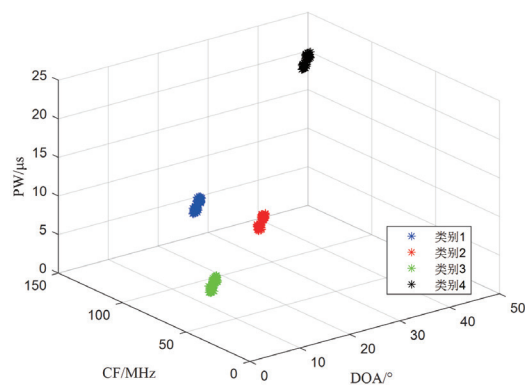


图3 动态 K-means 分选效果分布图

结果表明，动态 K-means 算法不需要预先监督学习即可将混合输入的信号进行分类识别出来。

3.2 SDIF 分选仿真

信号输入参数设置如表 2 所示，设置 8 个不同类型、参数设置的辐射源混合，经过动态 K-means 预分选后，将预分选结果输入到 SDIF 分选中，得到的混合辐射源分选结果如表 3 所示，仿真结果表明，SDIF 分选算法输出的辐射源参数正确，本算法能够正确分选出多个混合辐射源信号。

表 2 混合辐射源参数设置表

辐射源序号	AOA/(°)	CF/MHz	PW/μs	PRI/μs
1	24	200	20	80
2	10	300	5	40
3	15	360	500	6 000
4	35	390	0.3	30
5	50	400	5	6
6	5	430	70	200
7	39	470	50	600
8	26	300	12	95

表 3 混合辐射源仿真测试结果

辐射源序号	AOA/(°)	CF/MHz	PW/μs	PRI/μs
1	24	200.01	19.99	80
2	10	300.01	5	40.01
3	15	360.1	500	6000
4	35	390.1	0.3	30
5	50	400	5	5.99
6	5	430	70	199.97
7	39	470	49.99	600.01
8	26	300	12	95.03

4 结论

本文针对辐射源信号分选问题，结合信号脉内及脉间调制特征，提出一种动态 K-means 与 SDIF 相结合的分选方法。

该算法首先采用动态 K-means 算法对输入的脉冲信号进行预分选，稀释全脉冲序列中信号的密度，以此来减轻主分选的负担；后续主分选采用 SDIF 算法，通过计算相邻脉冲的 TOA 差构建第一级差值直方图，再通过序列搜索，分选出所有可能的辐射源脉冲序列。仿真结果表明，该算法具有较高的分选精度和鲁棒性以及较低的运算量，适合工程化实现。

参考文献：

[1] MARDIA H K. New techniques for the deinterleaving of repetitive of radar pulses[J]. Radar and signal processing. 1989, 136(4):149-154.

[2] MILOJEVIC D J, POPOVIC B M. Improved algorithm for the deinterleaving of radar pulses[J]. Radar and siganl processing, 1992,139(1): 98-104.

[3] 刘建新. 地波超视距模型下的时差提取及定位研究 [D]. 哈尔滨: 哈尔滨工业大学, 2019.

[4] 姜勤波, 马红光, 杨利锋. 脉冲重复间隔估计与去交织的正弦波插值算法 [J]. 电子与信息学报, 2007(2):350-354.

[5] 向炯, 汤建龙. 一种基于网格密度聚类的雷达信号分选 [J]. 火控雷达技术, 2010,39(4):67-72.

[6] 陈龙. 雷达信号分选跟踪器的设计与实现 [D]. 哈尔滨: 哈尔滨工程大学, 2011.

[7] 何川. 雷达信号重频分选算法概述 [J]. 电光系统, 2006,1: 4-6.

[8] 王大海. 雷达辐射源信号分选算法及硬件实现研究 [D]. 哈尔滨: 哈尔滨工程大学, 2014.

[9] 柴娟芳. 复杂环境下雷达信号的分选识别技术研究 [D]. 哈尔滨: 哈尔滨工程大学, 2009.

[10] 王石记, 司锡才. 雷达信号分选新算法研究 [J]. 系统工程与电子技术, 2003(9):1079-1083.

[11] 赵长虹. 重频分选与跟踪算法的研究 [D]. 西安: 西安电子科技大学, 2003.

[12] 刘明. 一种改进的基于 PRI 参数的雷达信号分选算法 [J]. 舰船电子工程, 2018,38(2):65-68.

[13] 王天航. 雷达信号分选及 DSP 实现 [D]. 哈尔滨: 哈尔滨工程大学, 2018.

【作者简介】

夏亦佳(1991—)，女，重庆人，硕士，工程师，研究方向：信号与信息处理。

(收稿日期：2024-12-19 修回日期：2025-04-28)