

# 一种新的未知雷达辐射源聚类分选方法

张红昌, 阮怀林, 龚亮亮, 罗景青

ZHANG Hong-chang, RUAN Huai-lin, GONG Liang-liang, LUO Jing-qing

电子工程学院, 合肥 230037

Electronic Engineering Institute, Hefei 230037, China

ZHANG Hong-chang, RUAN Huai-lin, GONG Liang-liang, et al. New clustering approach for sorting unknown radar emitter signal. Computer Engineering and Applications, 2008, 44(27): 200-202.

**Abstract:** Unknown radar signal sorting is always a problem of radar countermeasure. So the paper proposes a new signal sorting method based on data potential field. The method can cluster signals based on the relevance of signal parameters and does not need to appoint cluster number in advance. This method can process mass of data without any prior knowledge and is not sensitive to the influence of noise. In the end computer simulation results show that the proposed approach can feasibly and effectively sort the radar emitter signals.

**Key words:** radar signal sorting; cluster; data potential field; Electronic Warfare (EW)

**摘要:** 未知雷达信号的分选一直是雷达对抗情报处理中的难题。因此, 提出了一种新的基于数据势场聚类的未知雷达信号分选方法。该方法无需预先指定分类数, 利用信号参数之间的关联性自动进行聚类, 以此达到分选的效果。同时方法适宜于处理大量数据, 对噪声影响不敏感, 无需先验知识支撑。最后, 通过计算机仿真验证了该方法的可行性和有效性。

**关键词:** 信号分选; 聚类; 数据势场; 电子战

DOI: 10.3778/j.issn.1002-8331.2008.27.064 文章编号: 1002-8331(2008)27-0200-03 文献标识码: A 中图分类号: TN971.1

在现代电子战中, 雷达辐射源信号分选是电子侦察系统(ELINT)和电子支援系统(ESM)的重要环节之一, 是衡量雷达对抗情报处理能力的关键因素<sup>[1]</sup>。随着电子技术的飞速发展, 雷达的体制和信号调制样式变得越来越复杂和多样化。这样复杂的信号环境对分选方法的研究提出了更高的要求。对已知雷达的信号分选, 由于可以利用先验知识, 其处理的准确性和实时性已经达到了一定的高度。但对于未知雷达信号的分选则一直是分选算法研究中的难题。

现有的未知雷达信号分选方法<sup>[2-3]</sup>主要是基于信号子空间划分法。该方法将信号的特征描述参数张成的空间划分为多个分选子空间, 然后将脉冲描述字(PDW)在子空间上投影, 属于相同子空间的脉冲聚合在一起, 从而实现雷达信号分选。这类方法对常规体制雷达信号可以有效分选, 但对于复杂体制雷达或未知雷达的信号分选则面临着严峻的挑战。因此, 学者们在未知雷达信号分选研究中作了大量的探索, 提出了很多新思路和新方法。文献[4]中尝试利用 Kohonen 网络模型对雷达信号进行分选, 并对参数权值进行了确定, 提高了分选的准确率, 但算法中面临着对初始聚类中心数进行优化的问题。文献[5]中采用了 K-Means 聚类方法对雷达信号进行分选, 取得了较好的结果, 但该方法应用中仍需人为的指定分类数。因此, 本文提出了一种新的基于数据势场聚类的未知雷达信号分选方法, 该方法无需预先指定分类数, 利用信号参数之间的关联性自动进行聚类, 以此达到分选的效果。同时方法适宜于处理大量数据, 对噪

声影响不敏感, 无需先验知识支撑。

## 1 数据场聚类方法

### 1.1 数据势场

利用物理学中场和势的概念, 引入了数据场的势函数。通过势函数可以描述数据的空间分布, 利用所获的空间分布特性可实现对数据的聚类。定义两种形式势函数<sup>[6]</sup>: 拟重力场的势函数和拟核力场的势函数, 其形式如下。

拟重力场的势函数:

$$\varphi_x(y) = \frac{m}{1 + \left(\frac{\|x-y\|}{\sigma}\right)^k} \quad (1)$$

拟核力场的势函数:

$$\varphi_x(y) = m \times e^{-\left(\frac{\|x-y\|}{\sigma}\right)^k} \quad (2)$$

其中,  $m \geq 0$  代表场源强度, 可以看作数据对象的质量;  $\sigma \in (0, +\infty)$  用于控制对象间的相互作用力程, 称为影响因子;  $k \in N$  为距离指数。通过势函数可以对数据势场的空间分布性质进行描述。势函数实质上是位置或距离的函数, 可以叠加。因此, 数域空间中每一个数据对象对场中任何一点的势都有贡献, 且贡献的大小与两者间的距离平方成反比。直观上, 数据密集区应有较大的势, 而数据稀疏区的势也相对较小。可以采用物理学中等值线(面)来表述数据场的分布。

**作者简介:** 张红昌(1980-), 男, 硕士, 主研方向: 信号与信息处理; 阮怀林(1964-), 博士, 主研方向: 信号与信息处理。

**收稿日期:** 2008-03-21 **修回日期:** 2008-07-17

## 1.2 数据势场聚类

给定势值  $\Psi$ ,可在数据场中绘制一个相应的等势面,即满足  $\varphi(x)=\Psi$  的空间曲线,通过选定一组势值  $\{\Psi_1, \Psi_2, \dots\}$ ,即可用一系列的等势线来描述势函数在空间的分布。利用数据等势面的描述可实现对数据的聚类。将每个等势线(面)所包含的数据对象视为一个自然聚类,将不同等势线组成的嵌套结构视为类谱系图来实现不同层次的聚类划分。如图 1 所示为二维空间中随机取定 35 个等质量点,由拟核力场势函数产生的等势线分布示意图。

图 1、图 2 中显示:对应一组合合适取的势值,相应的等势线以不同的数据密集区为中心呈现自然的抱团特性。当势值  $\Psi=1.3687$  时,所有数据在层次 4 上聚为一类;随着势值的降低,在层次 3 上数据分为 ABC 和 DE 两类;继续降低后,数据将会不断的分类,最后当  $\Psi=5.4748$  时,数据分为本身的 5 类即 A、B、C、D、E 类。

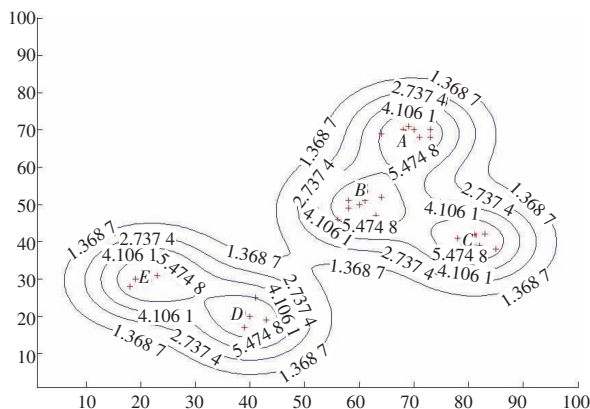


图 1 二维数据势场等势线分布

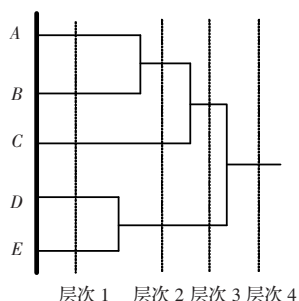


图 2 二维数据势场类谱图

## 1.3 影响因子确定

在采用势函数计算势场时,需要对式中的影响因子  $\sigma$  进行确定。如何在大量无先验知识的数据中确定适宜的影响因子对计算有着重要影响。影响因子决定了数据对象之间作用力程的影响半径,作用力程很短,则每个对象之间影响很小;作用力程太大,则相聚很远的也会相互作用。所以在影响因子  $\sigma$  的选取中应使势场分布尽可能的体现数据的内在分布。熵能够衡量数据分布的确定性和随机性。因而可以利用势熵来确定影响因子  $\sigma$ 。

令对象  $x_1, x_2, \dots, x_n$  的势值为  $\Psi_1, \Psi_2, \dots, \Psi_n$ , 势熵则可以如下定义<sup>[7]</sup>:

$$H = - \sum_{i=1}^n \frac{\Psi_i}{Z} \log\left(\frac{\Psi_i}{Z}\right) \quad (3)$$

式中,  $Z = \sum_{i=1}^n \Psi_i$  为标准化因子。

因此,当数据分布最大不确定时,熵最大,同样当数据的势值分布不对称,则不确定性最小,具有最小的熵。对应与影响因子则时,当  $\sigma \rightarrow 0$  时,每个数据均为单个作用的随机数,势熵最大;当  $\sigma \rightarrow \infty$  时,所有数据对应为一个随机数,势熵也最大;当  $\sigma$  为某个恰当值时,势熵最小,此时数据势场分布不均匀,不确定性最小。因此,可以通过计算势熵最小来获得优化影响因子。

## 2 未知雷达辐射源信号聚类分选算法

在雷达信号分选处理中,首先需要重叠的雷达信号进行初步的方位预选。方位信息是可靠的预选参数,只要雷达对抗侦察设备的方位分辨率足够高,可以将单位角度中的雷达信号密度降至很低。在此基础上在进行分选可以有效的降低处理数据量。在聚类分选中,选取雷达信号参数中载频 RF 和脉冲宽度 PW 作为距离计算的参数。载频和脉宽均具有量纲,其数值相差较大,需要对这两种参数进行归一化处理。

令  $f_{\max}, f_{\min}, \tau_{\max}, \tau_{\min}$  分别为带分选脉冲中载频最大值、载频最小值、脉宽最大值、脉宽最小值。则归一化参数可以通过如下计算获得:

$$f' = \frac{f}{f_{\max} - f_{\min}} \quad (4)$$

$$\tau' = \frac{\tau}{\tau_{\max} - \tau_{\min}} \quad (5)$$

本文选取势熵来确定影响因子  $\sigma$  的选取。确定影响因子后,则可以利用势函数计算带分选雷达脉冲信号分布所产生的势场分布,进而利用势场分布进行信号的聚类。本文提出了一种数据点近似势场的聚类处理方法其步骤如下:

(1) 利用数据势场分布,确定势场门限值  $\eta_0$ ;

(2) 剔除所有势值小于门限的数据点,获得大于门限值的的数据点集合  $\Omega$ ;

(3) 在  $\Omega$  中随机选择一个无类及处理标记的数据点  $x_i$ , 以  $x_i$  为中心寻找半径为  $\sigma' = \frac{3}{\sqrt{2}}\sigma$  的区域中所包含的其它数据点,计算这些数据点共同对点  $x_i$  产生的势值  $\Psi_i$ , 若  $\Psi_i \geq \eta_0$  则这些数据点归为一类,将这些数据点标记为类  $S_i$ , 同时  $x_i$  标记为已处理数据,其它数据仍标记为未处理数据;

(4) 选择有类标志  $S_i$  但没有处理标志的数据点,继续重复步骤(3),直到所有此类数据点处理完。这些数据全部标记为  $S_i$  类;

(5) 如有无类标志和未处理标记的数据,则重复进行步骤(3),更新  $S_i$  直到数据  $\Omega$  中无类和未处理标记的数据为止。

## 3 仿真分析

在实际的预选处理中,首先可以根据雷达使用的频率分段来对信号进行稀释,并进行方位预选。在此基础上获得的待分选信号中所含有的雷达数量将会大大降低。本文选取了 5 部信号参数较为接近且体制复杂的雷达辐射源进行仿真验证。这 5 部雷达信号参数设置如下表 1 所示。设这 5 部雷达辐射源信号已经过频段稀释和方位预选。这些雷达辐射源已处于雷达天线的测向分辨率内,产生的方位信息极为接近或相同,利用方位信息已无法进行分辨。此时利用这些雷达辐射源的载频和脉宽参数进行基于势场的聚类仿真分选。仿真所获得的雷达

表1 待分选雷达信号参数设置表

雷达编号	载频特征	载频/MHz	脉宽特征	脉宽/ $\mu$ s	重频特征	重复间隔/ $\mu$ s	信号个数
1	捷变	2 250~2 450	捷变	1.0~1.1	抖动	1 082~1 766	100
2	捷变	2 250~2 450	捷变	1.1~1.2	参差	1 067, 1 241	100
3	捷变	2 750~2 850	捷变	1.2~1.3	参差	1 331, 1 589, 1 048	100
4	捷变	2 600~2 750	捷变	1.1~1.2	滑变	983~1 023	100
5	捷变	2 550~2 750	捷变	1.0~1.1	抖动	939~973	100

信号在载频和脉宽二维参数上的分布如图1所示。

从图中可以看出5部雷达信号在载频和脉宽二维参数的分布相当的杂乱,不同雷达信号之间相互重叠严重。仿真设置的这些雷达信号中,其重复间隔之间非常的接近,并且还在不断变化,因此对于这种情况利用传统的方法很难将其分开。在此尝试利用数据势场聚类的方法对其进行分选,首先利用前面所述方法对参数进行归一化,并且随机取样20%数据点计算数据势场,搜索最优影响因子 $\sigma$ 。基于本文的仿真数据,计算获得 $\sigma=0.0012$ 。进而可以利用获取的优化影响因子对所有数据的场势分布进行计算获得如图3所示的场势分布图。

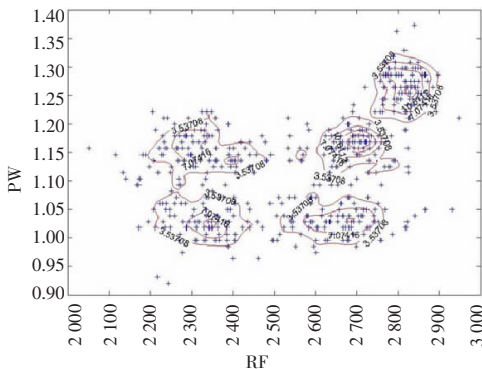


图3 待分选雷达信号数据势场等势线分布图

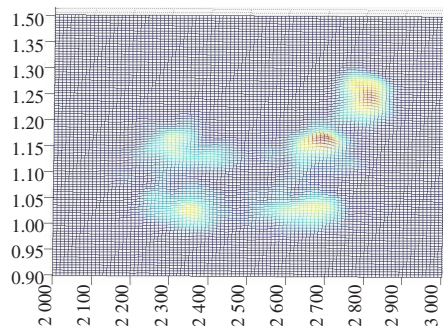


图4 数据势场分布3D图俯视图

由图3所示的等高势线分布情况可以得出当势值大于 $\eta_0=3.537$ 时,5部雷达的信号可以清楚的被分辨出。在图4中也可以看出数据势场中具有5个较为明显的势值峰。在此选用门限值,利用本文提出的基于数据点近似场势的聚类方法将不

同类的数据分离并聚类。经过计算可以获得如表2及图3所示的结果。通过公式(6)可求得总分选准确率为83.2%。

$$\text{分选准确率} = \frac{\text{准确分选的总信号数}}{\text{信号总数}} \times 100\% \quad (6)$$

由图3所示可以看出:雷达信号在产生时其相互重叠已经非常严重了,并且单个雷达的信号也是相当离散。但是在本文的分选处理中,在没有任何有关雷达信号先验知识的基础上,却仍然能够以较高的分选准确率将信号分选出来。说明此种方法在对未知复杂体制雷达信号的分选中是有效可行的。表2中可以看出,某些信号交叠比较严重的雷达,如雷达5有部分信号产生了误分选,信号交叠不严重时则有很好的结果如雷达1、2、4。同时由于算法中阈值的限定,部分离散数据会被作为噪声或干扰而被剔除除去,导致部分信号漏选,如雷达4、5。

表2 雷达信号聚类分选结果

雷达序号	1	2	3	4	5
实际脉冲数	100	100	100	100	100
分选脉冲数	89	92	103	78	88
正确分选脉冲数	86	88	91	77	74
漏分选脉冲数	14	12	9	23	26
误分选脉冲数	3	4	12	1	14

## 4 结论

基于数据势场的聚类方法具有层次聚类和密度聚类两种方法的优点,可以得到不同粒度上的多层次聚类结构,可以对任意形状的数据进行聚类并且对噪声数据不敏感。算法适宜于处理大量数据,并且无需任何先验知识的支撑。在复杂雷达信号环境的情况下,该方法对未知复杂体制雷达信号的分选中是可行的和有效的。采用聚类分选的方法会不可避免的导致部分信号的误分选,这也是聚类方法的通病。同时由于算法中阈值的限定会导致信号的漏选。因此如何动态选取阈值,减少误选和漏选数据数量,有待进一步研究。

## 参考文献:

- [1] 雷琴,金国庆,朱强华,等.复杂电磁环境下的雷达信号分选技术[J].舰船电子对抗,2007,30(4):74-76.
- [2] 李合生,韩宇,蔡英武,等.雷达信号分选关键技术研究综述[J].系统工程与电子技术,2005,27(12):2035-2040.
- [3] 赵国庆.雷达对抗原理[M].西安:西安电子科技大学出版社,1999.
- [4] 郭杰,陈军文.一种处理未知雷达信号的聚类分选方法[J].系统工程与电子技术,2006,28(6):853-856.
- [5] 张万军,樊甫华,谭营.聚类方法在雷达信号分选中的应用[J].雷达科学与技术,2004,2(4):219-223.
- [6] 戴晓军,凌文燕,李德毅.基于数据场的图像数据挖掘研究[J].计算机工程与应用,2004,40(26):41-43.
- [7] 李德毅,杜鹁.不确定性人工智能[M].北京:国防工业出版社,2005.