

# 基于核密度估计密度峰聚类的通信辐射源个体识别

李 昕, 雷迎科

(国防科技大学电子对抗学院, 合肥, 230037)

**摘要** 通信辐射源个体识别是通过发射机反映在信号上的差异来判别信号与辐射源个体之间的关联。传统的通信辐射源个体识别方法以及新兴的利用神经网络进行辐射源个体识别的方法都依赖带类别信息的信号样本,然而在实际中带类别信息的信号样本获取难度很大。为了解决这个问题,引入了无监督学习中的密度峰值聚类算法,在无类别信息信号样本的前提下进行通信辐射源个体识别。由于密度峰值聚类算法的性能受人工输入参数  $d_c$  的影响较大,文中利用核密度估计(KDE)及热扩散方程改进算法,在不需要人工输入参数的条件下实现对数据的分类。文中所提算法在实际电台信号数据集上进行了实验,具有较好的效果,验证了该算法的可靠性和有效性。

**关键词** 通信辐射源;个体识别;核密度估计;热扩散方程;聚类

**DOI** 10.3969/j.issn.1009-3516.2020.03.011

**中图分类号** TN975 **文献标志码** A **文章编号** 1009-3516(2020)03-0063-07

## Individual Communication Transmitter Identification Based on Density Peaks Clustering Algorithm via KDE

LI Xin, LEI Yingke

(College of Electronic Countermeasure, National University of Defense Technology, Hefei 230037, China)

**Abstract** Specific emitter identification (SEI) technique means that the relationship between the signal and the individual of the radiation source is judged by the difference in the signal of the transmitter response. Both the traditional methods and new emerging methods to make the neural networks for SEI rely on signal samples with category information. However, in conventional practice, the signal samples with category information are difficult to acquire. In order to solve this problem, this paper introduces a density peak clustering (DPC) algorithm in unsupervised learning to achieve SEI without classless information signal samples. Since the performance of the DPC algorithm is greatly influenced by the artificial input parameter  $d_c$ , this paper utilizes the diffusion equation and the kernel density estimation improved algorithm for classifying the data without the need of manual input parameters. The algorithm proposed in this paper is good in effects, and reliable and effectiveness in algorithm.

**Key words** specific emitter; identification; kernel density estimation; diffusion equation; clustering

通信辐射源个体识别是通过信号携带能够反映 发射机差异的信息及特征,来判别信号与发射机之

**收稿日期:** 2019-04-26

**作者简介:** 李 昕(1996—),男,安徽安庆人,硕士生,主要从事通信信号处理研究。E-mail:1515210772@qq.com

**引用格式:** 李昕,雷迎科. 基于核密度估计密度峰聚类的通信辐射源个体识别[J]. 空军工程大学学报(自然科学版), 2020, 21(3): 63-69. LI Xin, LEI Yingke. Individual Communication Transmitter Identification Based on Density Peaks Clustering Algorithm via KDE[J]. Journal of Air Force Engineering University (Natural Science Edition), 2020, 21(3): 63-69.

间的从属关系的,即通过信号区分发射机<sup>[1]</sup>。通信辐射源个体识别最早应用于军事通信领域<sup>[2]</sup>,并取得了较大的发展;最近随着新技术的发展,它在认知无线电<sup>[3]</sup>及自组织网络中扮演着越来越重要的作用<sup>[4]</sup>。

基于发射机的工作模式,通信辐射源个体识别使用的信号被分为瞬态信号和稳态信号。瞬态信号一般被称为“turn-on”信号,能够提供适合特征提取及发射器识别的独特且可区分的特性<sup>[5-6]</sup>。为了测量瞬态特征,主要方法是在噪声中检测瞬态信号的起点和终点<sup>[7]</sup>并提取它。瞬态特征基本上与发射机是一致的,这有利于识别;但瞬态信号持续时间较短,难以提取,且瞬态特征易受到非理想和复杂信道的影响,会造成识别结果不准确<sup>[8]</sup>。

稳态信号是发射机处于稳定工作状态下发射的信号,尽管稳态特征倾向于被传输信息的破坏程度,但对稳态信号的研究仍然有实际意义,因为稳态信号容易被检测和捕捉。在这种条件下很多特征提取方案已经被提出,例如双谱特征<sup>[9]</sup>和生物启发特征<sup>[10]</sup>。在这些方案中,使用的最为广泛的是基于信号时频表示的特征,时频表示是将信号映射到时间和频率的二维平面上,同时提供时间和光谱信息。文献<sup>[11]</sup>提出了一种基于短时傅里叶变换(STFT)的信号检测和识别系统,然而短时傅里叶系统是一种线性变换,不能用于分析非线性信号。文献<sup>[12]</sup>利用 Winger 和 Choi-William 分布提出了类似雷达波形识别算法。文献<sup>[13]</sup>使用了基于二次时频表示和顺序分类器的辐射源识别方案。XU<sup>[14]</sup>于2007年提出基于信号杂散特征识别电台的方法,可以识别同类型的电台个体,但在低信噪比条件下识别率不高。唐智灵<sup>[15]</sup>在2013年对杂散特征进行研究,在实际信号的数据集上验证了算法性能的可靠性。梁江海<sup>[16]</sup>利用经验模态分解模型,从时域和频域分析电台信号进行识别。韩洁等研究人员<sup>[17]</sup>在2017年提出将信号转换成3D-Hibert能量谱,在一定程度上实现了基于少量样本对通信电台个体的识别。唐哲等研究人员<sup>[18]</sup>提出利用矩形积分双谱特征和基于最大相关熵的通信辐射源识别方法。黄健航等研究人员<sup>[19]</sup>提出在小样本条件下基于自编码网络的通信辐射源个体识别方法。

以上提出的通信辐射源个体识别方法都需要充裕的已知类别信息的信号样本,而在实际复杂的电磁环境中,人们很难获得充裕的已知类别信息的辐射源观测样本数据。即在实际复杂电磁环境下,常常面临的是大规模、没有类别信息的通信辐射源观测样本数据。如果直接采用上述提出的基于有类

别信息的充裕样本的监督机器学习方法或者半监督机器学习方法,其识别性能必将受到影响。本文采用无监督机器学习中的聚类河源方法进行通信辐射源个体识别。

2014年,Alex Rodriguez 和 Alessandro Laio 提出了基于密度峰值聚类(Density Peaks Clustering, DPC)的算法<sup>[20]</sup>。DPC算法可以在无监督条件下对数据进行聚类,且算法简单高效,具有较好的识别效果。但DPC算法的识别效果受给定的数据库和挑选的参数 $d_c$ (截断距离)影响。 $d_c$ 主要用来计算各个点的密度并识别每个类的边界点。对于不同的数据集,DPC使用不同的方法去估计密度且参数 $d_c$ 的取决于主观经验。为了克服这个问题,本文采用基于核密度估计的方法改善DPC算法。本文提出的方法是一种基于热扩散方程<sup>[21]</sup>用于估计给定数据集概率分布的非参数方法,该方法是基于在无限域中的热扩散,考虑到了截断距离的选择和核密度估计的边界校正。

本文提出的基于改进的密度峰值聚类的通信辐射源个体识别方法能够在无监督条件下进行通信辐射源个体识别,并解决已知类别信息的信号样本数量不足问题;针对DPC算法识别效果依赖于人工输入参数 $d_c$ ,在球形数据集上聚类性能不佳的问题,利用热扩散方程中时间参数 $t$ 计算各据点的密度和寻找各簇的边界,实现参数自适应,在不需要人工输入参数 $d_c$ 时在各个数据集上都有较好的表现。

## 1 DPC 算法

DPC算法能够通过快速搜索聚类中心的方法来创建任意形状的簇,该算法认为作为聚类中心的数据点具有较高的局部密度 $\rho_i$ ,与非聚类中心点相比,到其他聚类中心点的距离 $d$ 更大。因此对于数据点 $i$ ,DPC算法需要计算局部密度 $\rho_i$ 以及点 $i$ 到最近的且密度比其大的点的距离 $\delta_i$ 。具体步骤如下:

1)计算局部密度 $\rho_i$ 。

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \quad (1)$$

其中:

$$\chi(d) = \begin{cases} 1, & d < 0 \\ 0, & d \geq 0 \end{cases} \quad (2)$$

式中: $d_{ij}$ 表示点 $i$ 到点 $j$ 的距离; $d_c$ 为截断距离; $\rho_i$ 为距点 $i$ 距离小于 $d_c$ 点的数量。因此, $d_c$ 在计算每个点的局部密度时是必不可少的参数,DPC算法的性能也取决于参数 $d_c$ 的选取。对于小的数据集, $\rho_i$

可能会受到较大的统计误差的影响,因此本文采用非参数核密度估计方法<sup>[22]</sup>计算各个点的局部密度。

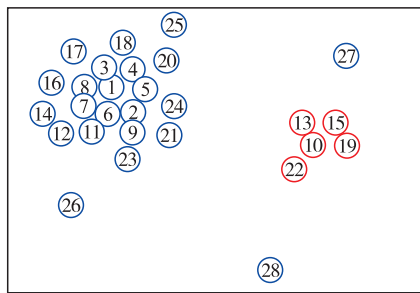
2)计算相对距离  $\delta_i$ 。

$$\delta_i = \begin{cases} \max(d_{ij}), & \rho_i > \rho_j \\ \min(d_{ij}), & \text{other} \end{cases} \quad (3)$$

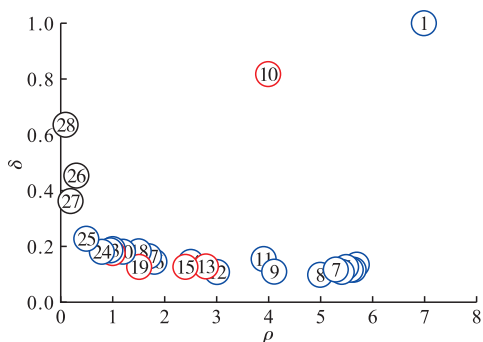
式中: $\delta_i$ 的定义是数据点  $i$  到其最近的且局部密度比其大的数据点的距离,如果数据点  $i$  是密度最大的点,则  $\delta_i$  为点  $i$  到其它点距离的最大值。

3)寻找聚类中心点。

聚类中心点具有较高的局部密度以及较大的相对距离,因此,计算完各个数据点的局部密度  $\rho_i$  和相对距离  $\delta_i$  后,以  $\rho$  和  $\delta$  为坐标轴建立二维平面直角坐标系,各个数据点根据各自的  $\rho_i$  值和  $\delta_i$  值放置在该坐标系中,形成二维平面图被称为决策图。如图 1(a)所示,28 个数据点按密度递减顺序排列,图 1(b)是相对应的这 28 个点形成的决策图。点 1 和 10 具有较高的  $\rho$  值和  $\delta$  值,这是聚类中心点的典型特征。因为点 26,27 和 28 是孤立的,它们具有高  $\delta$  和低  $\rho$ ,可以被认为是噪声或异常值。因此,使用决策图,可以很容易地识别聚类中心点。在成功找出聚类中心点后,DPC 会根据  $\delta$  值将剩余数据点分配到最近的聚类中心。



(a) 28 个示例数据点



(b) 对应点的决策图

图 1 决策图示例

4)寻找边界区域。

归属于一个簇的数据点也在另一个聚类中心点的距离小于  $d_c$ ,这类点的集合被称为该簇的边界区域。对于这些边界点,DPC 在的边界区域内找到具有最大密度点的密度为  $\rho_b$ ,密度高于  $\rho_b$  的那些点被

认为是聚类点,而其他数据点被识别为簇晕点或者噪音。

## 2 基于热传播的密度峰值聚类

### 2.1 核密度估计

非参数密度估计是统计分析数据的重要工具之一,它可用于评估偏度,多模态统计,贝叶斯后验总结,判别分析和分类<sup>[22-23]</sup>。与传统方法相比<sup>[22]</sup>,非参数方法对数据集建模更灵活,并且不受参数设定偏差的影响<sup>[23]</sup>。核密度估计是最常用的非参数密度估计方法。用于估计密度的现有技术方法包括引入一个窄带高斯核—— $\hat{f}h(i)$ 。对于每一个数据点  $i$ ,在整个数据集上计算所有内核值的积分<sup>[24-25]</sup>。用于进行相同且独立的未知概率密度函数(Probability Density Function, PDF)的数据点  $\{x_1, x_2, \dots, x_n\}$  的核密度估计为:

$$\hat{f}h(x_i; h) = \frac{1}{n} \sum_{j=1}^n K_h(x - x_j) \quad (4)$$

高斯核  $K(x; x_j; h)$  经常被用做估计密度。

$$K(x; x_j; h) = \frac{1}{\sqrt{2\pi}h} e^{-\frac{(x-x_j)^2}{2h^2}} \quad (5)$$

式中: $K$  是按照  $1/h$  缩放的核函数; $h$  是核函数的带宽。

式(4)的性能很大程度上依赖于对带宽  $h$  的选择<sup>[26-27]</sup>。积分均方误差(Mean Integrated Squared Error, MISE)<sup>[21]</sup>是用于确定  $h$  的最佳值的工具之一,如式(6)所示:

$$MISE\{\hat{f}\}(h) = E_f \int [\hat{f}(x; h) - f(x)]^2 dx \quad (6)$$

但是高斯核密度估计存在以下问题:①参数  $h$  (带宽)难以选择;②存在边界误差;③平滑不足或者过度平滑。

### 2.2 改进后的核密度估计

不同于式(1)和(4)的方法,本文采用热扩散方程<sup>[21]</sup>进行核密度估计。热扩散方法将核密度估计视为扩散偏微分方程的唯一解,其时间参数  $t$  与核带宽  $h$  意义一致<sup>[28-29]</sup>。通过热扩散方程对 KDE 的解释源于维纳过程——对于一个连续时间随机过程,其下一阶段可以由前一状态直接计算,具体解释如下所示:

- 1)  $n$  个均匀分布的  $l$  维数据点  $\{x_1, x_2, \dots, x_n\}$ ;
- 2) 采用高斯核计算从点  $x_i$  到  $x_j$  的转移概率:

$$P_t = \frac{1}{n} \sum_{j=1}^n \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-x_j)^2}{2t}} \quad (7)$$

核密度估计在式(7)中被解释为在该过程中与时间  $t$  有关的概率密度分布函数,参数  $t$  与式(5)中

的参数  $h$  一致。

$$\hat{f}(x;t) = \frac{1}{2} \sum_{j=1}^n \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-x_j)^2}{2t}} \quad (8)$$

式(8)是一个迭代过程,因此这个转变满足偏微分方程。

$$\frac{\partial}{\partial t} \hat{f}(x,t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \hat{f}(x,t), \quad x \in X, t > 0 \quad (9)$$

式中:  $X \equiv \mathbb{R}$ 。初始条件为  $\hat{f}(x;0) = \Delta(x)$ , 其中  $\Delta(x) = \frac{1}{n} \sum_{j=1}^n \delta(x-x_j)$ , 这是数据集  $X$  的经验密度。 $\delta(x-x_j)$  是狄拉克函数, 它将点质量分配给数据集上所有的数据点。当域具有有限的端点时, 式(4)需要边界校正。因此, 以偏微分方程为基础, 利用 Neumann 边界条件(见式(10))和初始条件  $\hat{f}(x;0) = \Delta(x)$  求解式(9)。

$$\frac{\partial}{\partial t} \hat{f}(x;t) \Big|_{x=X_l} = \frac{\partial}{\partial t} \hat{f}(x;t) \Big|_{x=X_u} = 0 \quad (10)$$

式中:  $X_l$  和  $X_u$  是域对应的上限和下限。考虑到 Neumann 边界条件和数据点的概率密度在  $[0, 1]$  内, 因此这个偏微分方程的解可以用  $\theta$  核( $\theta$ )来代替高斯核表示。

$$\hat{f}(x;t) = \frac{1}{n} \sum_{j=1}^n \theta(x; x_j; t), \quad x \in [0, 1] \quad (11)$$

其中:

$$\theta(x; x_j; t) = \sum_{k=-\infty}^{+\infty} [\varphi(x, 2k + x_j; t) + \varphi(x, 2k - x_j; t)] \quad (12)$$

然后式(11)可以被写为:

$$\hat{f}(x;t) = \frac{1}{n} \sum_{j=1}^n \sum_{k=-\infty}^{+\infty} e^{-k^2 \pi^2 t/2} \cos(k\pi x) \cos(k\pi x_j) \quad (13)$$

则:

$$\hat{f}(x;t) \approx \sum_{k=0}^{n-1} a_k e^{-k^2 \pi^2 t/2} \cos(k\pi d) \quad (14)$$

$n$  是一个正整数, 则:

$$a_k = \begin{cases} 1, & k = 0 \\ \frac{1}{n} \sum_{i=1}^n \cos(k\pi - x_i), & k = 1, 2, \dots, n-1 \end{cases} \quad (15)$$

式(14)是核密度估计的自适应和替代形式, 并且考虑了最佳带宽选择和边界校正。此外, 式(14)可以用快速傅里叶变换求解<sup>[23]</sup>。带宽较小时, 式(13)与高斯核相似; 而带宽较大时, 是一个统一的内核<sup>[28]</sup>。因此它的性能更优越且与数据的真实密度一致, 而式(4)计算得到的密度与真实密度是不一致的。

### 2.3 最佳带宽自适应选择

Sheather-Jones(SJ)算法<sup>[21]</sup>曾被用来计算最佳

带宽, 本文采用改进后的 Sheather-Jones 算法(The Improved Sheather-Jones)来计算最佳带宽。递归的定点解被看作是带宽的最佳值, 并且可以使用快速余弦变换来估计, 而不需要考虑分布的正态性假设。利用非线性方程的唯一解可以自适应找到核密度估计的最佳带宽  $t$ 。

$$t = \xi \gamma^{[\cdot]}(t) \quad (16)$$

SJ 算法利用式(17)计算带宽,  $*\hat{t}_{l+1}$  通过式(18)估计。这个假设的前提是  $f$  是正态分布密度且可以估算出数据的均值和方差。这个迭代过程的最大弱点是假设  $f$  为正态密度, 这个假设会导致  $*\hat{t}$  出现随机错误, 例如  $f$  与高斯分布相差较大时。因此本文利用非线性方程求解以获得带宽的最佳值。

$$*\hat{t} = \epsilon \gamma^l(*\hat{t}_{l+1}) \quad (17)$$

$$*t_j =$$

$$\left( \frac{1+1/2^{(j+1)/2}}{3} \cdot \frac{1 \times 3 \times 5 \times \dots \times (2j-1)}{N \sqrt{\pi/2} \|f^{(j+1)}\|^2} \right)^{2/(3+2j)} \quad (18)$$

ISJ 算法的具体流程为:

1) 初始化  $z_0 = \epsilon$ ,  $\epsilon$  是计算精度,  $n=0$ ;

2) 计算  $z_{n+1} = \xi \gamma^{[l]}(z_n)$  (一般  $l$  的取值为 5);

3) 如果  $|z_{n+1} - z_n| < \epsilon$ , 迭代停止且  $*t = z_{n+1}$ ; 否则  $n = n+1$ , 重复步骤 2);

4) 在  $*t$  处评估的高斯核密度估计量作为  $f$  的最终估计量;  $*t_2 = \gamma^{[l-1]}(z_{n+1})$  作为  $\|f''\|$  的最佳估计带宽。

最佳带宽  $t$  对核函数进行规范以获得更准确的估计密度, 本文采用  $t = \text{sqrt}(t)/3.3$  去细化簇的边界。

## 2.4 本文进行通信辐射源个体识别算法具体步骤

**步骤 1** 对信号进行矩形积分双谱变换<sup>[14]</sup>然后计算直方图特征<sup>[30]</sup>;

**步骤 2** 计算各个信号直方图特征的欧式距离;

**步骤 3** 通过式(16)计算带宽  $t$ ;

**步骤 4** 通过式(1)计算  $\rho_i$ , 通过式(3)计算  $\delta_i$ ;

**步骤 5** 画出决策图并选择聚类中心;

**步骤 6** 给其余非中心点分类, 并检查簇的边界点条件;

**步骤 7** 输出各信号的类别。

## 3 实验与分析

### 3.1 算法在机器学习数据集上的表现

为了验证采用模糊邻域关系比较密度峰值算法的有效性, 本文将其与 DBSCN 算法<sup>[31]</sup>、OPTICS 算

法<sup>[32]</sup>以及 DPC 算法<sup>[20]</sup>进行了比较。实验数据集有源于文献<sup>[31]</sup>中的 Aggregation 数据集、D31 数据集和 R15 数据集以及来自于 UCI 机器学习库中的 OptdigitsDataset (Optdigits)、Sonar Dataset (Sonar)、the Heart Disease database (Heart)、Wine Recognition Database (Wine) 和 Iris Plants Database(Iris)。上述数据集的具体说明见表 1。

表 1 实验数据集描述

数据集	数据数量	数据维度	数据类别数
D31	3 100	2	31
Optdigits	5 620	64	10
Aggregation	788	2	6
R15	7 400	2	15
Sonar	208	60	2
Heart	270	13	2
Wine	178	13	3
Iris	150	4	3

表 2 列出了本文提出的算法和其它算法在 8 个数据集上的准确率。对于 Heart 这个数据集, DBSCN 算法出现了异常值,因此没有在这个数据集上对 DBSCN 算法进行评估。从中我们可以看出, DPC-KDE 算法能更准确地识别聚类中心点,而不依赖于数据集的性质,而 DPC 查找密度和边界点的能力在很大程度上取决于数据集的性质。本文提出的 DPC-KDE 方法是不同数据集聚类的有效通用解决方案。DPC 算法的性能很大程度上取决于参数  $d_c$  的选取,而  $d_c$  的选取很大程度上依赖于主观经验,这对算法性能的提升是很大的制约,本文提出 DPC-KDE 算法能够自适应选择最佳参数,在任一数据集上都有较好的表现。

表 2 本文算法与其它算法在不同数据集上的对比

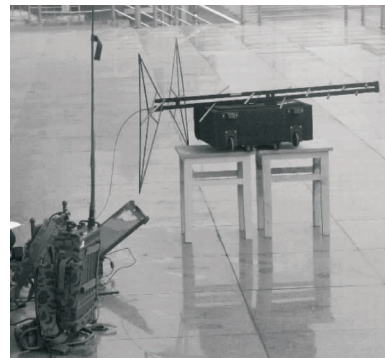
数据集	DBSCN	OPTICS	DPC	DPC-KDE
D31	0.895	0.895	0.912	1.00
Optdigits	0.58	0.58	0.6	0.72
Aggregation	0.9	0.9	0.903	0.997
R15	0.952	0.952	0.96	1.00
Sonar	0.612 5	0.612 5	0.68	0.78
Heart		0.78	0.804 7	0.862
Wine	0.339 2	0.339 2	0.915 7	0.956
Iris	0.667	0.667	0.9	0.973

### 3.2 算法在电台信号数据集上的表现

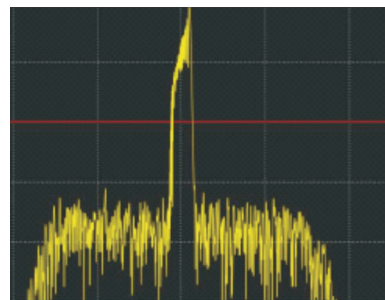
#### 3.2.1 电台信号数据集

为了验证本文提出识别方法的可行性,算法在实际超短波(USW)数据集进行了实验,数据采

集场景见图 2。课题组采集了 5 部超短波 FM 电台不同说话人的语音通信数据。超短波电台信号数据采集时,使用手机播放 3 个不同的人说话声音作为发射端的输入,然后接收机作为非协作方经过采集获得零中频 I/Q 信号。背负式超短波 FM 电台中心频率分别设置为 35 MHz、55 MHz 和 85 MHz,工作模式为“小功率”,并设立了 3 个采集场景:接收机与电台之间无高大建筑遮挡且距离分别为 100 m 和 50 m,以及接收机与电台中间存在高大建筑物遮挡且距离为 50 m,其余参数见表 3。信号采集完之后,以 2 048 个点为单位进行分割,然后分别计算每段信号的直方图特征(特征的维度可以调整),最后将特征放入算法中进行计算并识别各段信号。



(a) USW 数据集采集实况



(b) 电台 1 信号采集界面频域

图 2 数据采集场景

表 3 USW 数据集采集接收机参数

信号带宽/kHz	接收机采集增益/dB	接收机信道带宽/kHz	采样频率/kHz	采样时间/s
11.2	6	100	312.5	180

#### 3.2.2 特征维度对识别效果的影响

如表 4 所示,在不同直方图特征维度下进行实验,同其他方法相比,本文方法表现出优异的性能。此外,当直方图特征的维数从 128 增加到 256 时,所有方法的识别率都得到显著改善。可以推断出,不同的维度直方图特征包含的信息量不同。特征的维度越高,它们包含的能够区分不同电台信号的信息就越多。

表4 不同特征维度各算法的识别率

Dimensions	$k$ -means	DBSCN	DPC	DPC-KDE
32	32.7%	32.9%	34.2%	35.1%
64	32.7%	33.4%	34.7%	35.6%
128	33.5%	34.2%	35.8%	36.4%
256	64.2%	65.6%	69.1%	74.6%
512	64.2%	66.3%	69.1%	74.7%

### 3.2.3 信号信噪比(SNR)对识别效果的影响

表5是DPC-KDE, DPC, DBSCN和 $k$ -means算法在不同信噪比时的识别率。结果表明,在不同信噪比下, DPC-KDE在识别率方面优于其他聚类方法。它表明信号的信噪比越高,发射机反应在信号的差异越大。

表5 不同信噪比各算法识别率

SNR/dB	$k$ -means	DBSCN	DPC	DPC-KDE
0	50.2%	45.7%	54.2%	54.6%
10	55.8%	53.4%	60.1%	63.4%
20	60.2%	61.5%	64.9%	70.0%
30	64.2%	66.3%	69.1%	74.7%

## 4 结语

本文将密度峰值算法引入到通信辐射源个体识别中,解决以往全监督或半监督方法在缺乏足够带类别信息的信号样本表现不佳问题,在无监督条件下实现通信辐射源个体识别;针对DPC算法的性能受数据集特性的影响,提出利用核密度估计对数据的内核自适应建模,在不同的数据集上有不同的密度计算方式不同,使估算出的密度更逼近于数据的真实密度;针对DPC算法的参数 $d_c$ 的取值一般由主观经验决定,影响了算法的性能,提出利用热扩散方程中的时间参数 $t$ 代替带宽 $h$ ,然后采用ISE算法自适应获得参数的最佳值,进一步提升算法的性能。本文的提出DPC-KDE在实际电台数据集上得到了验证,具有较好的识别效果,在缺乏甚至没有带类别信息的信号样本时,能够发挥重要作用。但本文的方法与以往方法一样,都是在闭集上进行识别,影响其实用性,接下来要结合增量学习,解决通信辐射源个体识别模型的在线学习问题。

### 参考文献

[1] HUANG G, YUAN Y, WANG X, et al. Specific Emitter Identification for Communications Transmitter Using Multi-Measurements [J]. Wireless Personal Communications, 2017, 94(3):1523-1542.

[2] WILEY R G. Electronic Intelligence: The Interception of Radar Signals[J]. Electronics & Power, 1986, 32(11,12):829.

[3] SAMUEL J N, DU PLESSIS W P. Specific Emitter Identification for Enhanced Access Control Security[J]. SAIEE Africa Research Journal, 2017, 108(2):71-79.

[4] ZHANG Z, LONG K, WANG J. Self-Organization Paradigms and Optimization Approaches for Cognitive Radio Technologies: a Survey [J]. IEEE Wireless Communications, 2013, 20(2):36-42.

[5] ZHU S L, GAN L. Specific Emitter Identification Based on Visibility Graph Entropy[J]. Chinese Physics Letters, 2018, 35(3):9-13.

[6] URETEN O, SERINKEN N. Bayesian Detection of WiFi Transmitter RF Fingerprints [J]. Electronics Letters, 2005, 41(6):373.

[7] HUANG Y, DJURIC P M. Bayesian Detection of Transient Signals Incolored Noise[C]//2000 IEEE International Conference on Acoustics, Speech and Signal Processing. Istanbul, Turkey: IEEE, 2000: 725-728.

[8] 徐书华. 基于信号指纹的通信辐射源个体识别技术研究[D]. 武汉:华中科技大学, 2007.

[9] 刘莹, 单洪, 胡以华, 等. 基于双谱二次特征的通信信号识别算法研究[J]. 计算机应用研究, 2017, 34(6): 1835-1837.

[10] ZHANG Z, LONG K, WANG J, et al. On Swarm Intelligence Inspired Self-Organized Networking: Its Bionic Mechanisms, Designing Principles and Optimization Approaches[J]. IEEE Communications Surveys & Tutorials, 2014, 16(1):513-537.

[11] LOPEZ R G, GRAJAL J, SANZ O A. Digital Channelized Receiver Based on Time-Frequency Analysis for Signal Interception[J]. IEEE Transactions on Aerospace and Electronic Systems, 2005, 41(3):879-898.

[12] LI L, JI H B, JIANG L. Quadratic on Time-Frequency Analysis and Sequential Recognition for Specific Emitter Identification[J]. IET Signal Processing, 2011, 5(6):568-574.

[13] GILLESPIE B W, ATLAS L E. Optimizing Time-Frequency Kernels for Classification[J]. IEEE Transactions on Signal Processing, 2001, 49(3):485-496.

[14] XU S H. Radio Transmitter Classification Using a New Method of Stray Features Analysis Combined with PCA[C]//Military Communication Conference. Orlando, FL, USA: IEEE, 2007:1-5.

[15] 唐智灵. 通信辐射源非线性个体识别方法研究[D]. 西安:西安电子科技大学, 2013.

[16] 梁江海. 一种基于经验模态分解的通信辐射源个体识别方法[J]. 中国电子科学研究院学报, 2013, 8(4): 393-397.

- [17] 韩洁, 张涛, 王欢欢, 等. 基于 3D-Hibert 能量谱和多尺度分形特征的通信辐射源个体识别[J]. 通信学报, 2017, 38(4): 99-109.
- [18] 唐哲, 雷迎科. 基于最大相关熵的通信辐射源个体识别方法[J]. 通信学报, 2016, 37(12):171-175.
- [19] 黄健航, 雷迎科. 通信辐射源个体识别的自编码器构造方法[J]. 火力与指挥控制, 2018, 43(11):110-114.
- [20] RODRIGOEZ A, LAIO A. Clustering by Fast Research and Find of Density Peks[J]. Science, 2014, 344(6191):1492-1496.
- [21] BOTEV Z I, GROTEWSKI J F, KROESE D P. Kernel Density Estimation via Diffusion[J]. The Annals of Statistics, 2010, 38(5):2919-2957.
- [22] 孙志华. 非参数与半参数统计[M]. 北京:清华大学出版社, 2016.
- [23] LEHMANN E L. Model Specification; the Views of Fisher and Neyman, and Later Developments [M]. Boston, MA: Springer, 2012:955-963.
- [24] ROSENBLATT M. Remarks on Some Nonparametric Estimates of a Density Function[J]. The Annals of Mathematical Statistics, 1956, 27(3):832-837.
- [25] MARGOLIN A A, WANG K, LIM W K, et al. Reverse Engineering Cellular Networks[J]. Nature Protocols, 2006, 1(2):662-671.
- [26] JONES M C, MARRON J S, SHEATHER S J. A Brief Survey of Bandwidth Selection for Density Estimation[J]. Journal of the American Statistical Association, 1996, 91(433):401-407.
- [27] SHEATHER S J, JONES M C. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation[J]. Journal of the Royal Statistical Society Series B (Methodological), 1991, 53(3):683-690.
- [28] KRISHNASWAMY S, SPITZER M H, MINGUE-NEAU M, et al. Conditional Density-Based Analysis of T Cell Signaling in Single-Cell Data[J]. Science 2014, 346(6213):1250689.
- [29] XU X Y, YAN Z, XU S L. Estimating Wind Speed Probability Distribution by Diffusion-Based Kernel Density Method [J]. Electric Power Systems Research, 2015, 121:28-37.
- [30] 余旺盛, 李卫华, 侯志强. 分层结构直方图及其应用[J]. 电子学报, 2017, 45(11):2617-2624.
- [31] 秦佳睿, 徐蔚鸿, 马红华. 自适应局部半径的 DB-SCAN 聚类算法[J]. 小型微型计算机系统, 2018, 39(10):60-64.
- [32] 安建瑞, 张龙波, 王雷, 等. 一种基于网格与加权信息熵的 OPTICS 改进算法[J]. 计算机工程, 2017, 43(2):206-209.

(编辑:徐楠楠)

\*\*\*\*\*

## (上接第 30 页)

- [5] 刘玮, 宋贵宝, 陈小卫. 基于粗糙集和神经网络的导弹故障诊断方法[J]. 海军航空工程学院学报, 2009, 24(2):98-100, 104.
- [6] 朱松青, 史金飞. 状态监测与故障诊断中的主元分析法[J]. 机床与液压, 2007, 35(1):246-248.
- [7] DING S X. Model-Based Fault Diagnosis Techniques-Design Schemes, Algorithms and Tools [J]. IFAC Papers on Line, 2016, 49(15):50-56.
- [8] XUE X D, LIU F L, LIU D Y, et al. Design and Realization of Inference Engine of Fault Diagnosis Expert System for Electronic Recycling Equipment of Missile Based on Support Vector Machine [C]//The 2nd IEEE International Conference on Computer and Communications (ICCC). IEEE; [s. n.], 2016. doi: 10.1109/comppcomm.2016.7924.914.
- [9] YANG J B, LIU J, WANG J, et al. Belief Rule Base Inference Methodology Using the Evidential Reasoning Approach-RIMER [J]. IEEE Transactions on Systems Man & Cybernetics Part A Systems & Humans, 2006, 36(2):266-285.
- [10] XU D L, LIU J, YANG J B, et al. Inference and Learning Methodology of Belief-Rule-Based Expert System for Pipeline Leak Detection [J]. Expert Systems with Applications, 2007, 32(1):103-113.
- [11] 齐铎, 范惠林, 陈丹强, 等. 基于神经网络的红外型空空导弹故障诊断专家系统[J]. 计测技术, 2012, 32(3):24-26, 30.
- [12] 樊会涛, 崔颢, 天光. 空空导弹 70 年发展综述[J]. 航空兵器, 2016, 291(1):7-16.
- [13] GUO Q E, WANG X Q, WEI Z. Fuzzy Comprehensive Evaluation Based on Cross-evaluation and Its Application [J]. Control & Decision, 2012, 27(4):575-578, 583.
- [14] 梁杰, 陈嘉豪, 张雪芹, 等. 基于独热编码和卷积神经网络的异常检测[J]. 清华大学学报(自然科学版), 2019, 59(7):523-529.
- [15] 王三煜. 制冷器在导弹系统中的应用[J]. 红外技术, 2005, 27(5):59-62.
- [16] HON R C, KIRKCONNELL C S. Raytheon Dual-use Long Life Cryocooler Development [C]//Defense & Security Symposium. International Society for Optics and Photonics, 2007. http://doi: org/10.1117/12.719956.

(编辑:徐敏)