

BDmin-VMD-CA 结合 MDFF 的通信辐射源个体识别^①



刘高辉, 闫迪

(西安理工大学 自动化与信息工程学院, 西安 710048)

通信作者: 刘高辉, E-mail: liugh68@xaut.edu.cn

摘要: 针对通信辐射源细微指纹特征难以提取及单一特征识别率不高的问题, 提出了一种联合最小巴氏距离和相关性分析的变分模态分解与多域特征参数融合的通信辐射源个体识别方法. 首先, 采用基于最小巴氏距离的变分模态分解方法对通信辐射源信号的每个符号波形进行分解, 得到若干个包含数据信息的低频本征模态函数和包含指纹信息的高频本征模态函数; 然后, 计算各本征模态函数与其符号波形信号的相关系数, 选取相关系数小的本征模态函数作为辐射源细微特征分量, 对细微特征分量提取时域、频域及熵多特征参数, 并拼接融合成多域特征向量实现对通信辐射源符号波形的特征提取; 最后, 通过长短期记忆网络对辐射源信号每个符号的多域特征向量依次进行学习分类, 实现通信辐射源个体识别. 选择公开的 Oracle 数据集进行了实验验证, 实验结果表明当信噪比为 6 dB 时, 本文提出的方法识别准确率可达 96.7%, 比各单一域平均识别准确率提高了 22.1%.

关键词: 通信辐射源个体识别; 巴氏距离; 相关性分析; 变分模态分解; 多域特征融合

引用格式: 刘高辉, 闫迪. BDmin-VMD-CA 结合 MDFF 的通信辐射源个体识别. 计算机系统应用, 2025, 34(8): 105-115. <http://www.c-s-a.org.cn/1003-3254/9916.html>

Identification of Communication Radiation Source Individuals by BDmin-VMD-CA Combined with MDFF

LIU Gao-Hui, YAN Di

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: Addressing the challenges in extracting subtle fingerprint features of communication emitters and the low recognition rate of single-feature identification, a method for individual identification of communication emitters is proposed. This method combines variational mode decomposition (VMD) with multi-domain feature fusion (MDFF), based on the joint application of minimum Bhattacharyya distance (BDmin) and correlation analysis (CA). First, the VMD method based on BDmin is utilized to decompose each symbol waveform of the communication emitter signal into several intrinsic mode functions (IMFs), including low-frequency IMFs containing data information and high-frequency IMFs containing fingerprint information. Next, the correlation coefficients between each IMF and its symbol waveform signal are calculated, with IMFs exhibiting small correlation coefficients being selected as the subtle feature components of the emitter. Time-domain, frequency-domain, and entropy-based multi-feature parameters are then extracted from these subtle feature components and concatenated into a multi-domain feature vector for feature extraction of the communication emitter's symbol waveform. Finally, a long short-term memory (LSTM) network is employed to sequentially learn and classify the multi-domain feature vectors of each symbol of the emitter signal, thus achieving individual identification and classification of communication emitters. Experimental validation is conducted using the publicly available Oracle dataset. The results show that, when the signal-to-noise ratio (SNR) is 6 dB, the proposed

^① 收稿时间: 2024-12-31; 修改时间: 2025-02-12; 采用时间: 2025-02-18; csa 在线出版时间: 2025-06-13

CNKI 网络首发时间: 2025-06-16

algorithm achieves a recognition accuracy of 96.7%, representing a 22.1% improvement compared to the average recognition accuracy of individual domains.

Key words: identification of communication radiation source individual; Bhattacharyya distance (BD); correlation analysis (CA); variational mode decomposition (VMD); multi-domain feature fusion (MDFF)

通信辐射源个体识别技术是利用信号处理并结合分类算法和深度学习算法,通过提取辐射源信号所携带的细微指纹特征来识别特定发射器个体的技术^[1]。目前,通信辐射源个体识别技术已广泛应用于军用和民用领域,并因此成为近年来国内外学者的研究热点之一。

由于通信信号调制方式日益复杂化,仅依赖单一特征识别已经难以全面捕捉并反映信号间细微而关键的差异,越来越多的研究者对信号提取多特征来实现通信辐射源识别。文献[2]通过对 OFDM 前同步码经过多域信息处理后,再将其输入到 DRSN 网络中,对其进行特征提取与分类,再将原始分类结果进行整合融合,解决了单一特征表达能力不足的问题。文献[3]提取一维时域信号和二维时频分布信号的数据特征后,将融合特征输入到残差神经网络中分类识别,通过实测数据的分析结果表明,与传统的特征提取方法相比,该方法具有更高的精度。然而,缺点是缺乏对数据的非线性描述和非线性特征对识别精度的影响。随着神经网络的快速发展,深度学习在 SEI 中得到了广泛的应用。文献[4]通过将特征图叠加与神经网络模型相结合,比单一特征更能反映不同发射体的指纹特征。文献[5]将信号的双谱、周期谱和功率谱等作为传统特征,并利用残差神经网络从原始信号中提取固有特征,将残差神经网络提取的传统信号特征与固有特征融合实现了多特征融合。然而,深度学习模型的训练往往需要大量的训练样本。实际上,获取大量实际探测到的通信辐射源信号是具有挑战性的。

另外,由于通信辐射源个体细微特征具有非线性,而熵通常被视为描述数据或系统非线性特征的一种度量。在文献[6]中,作者采用了排列熵作为区分不同辐射源个体特征的方法。但是,该方法并未考虑信号幅度间的相互关系。在 2016 年, Rostaghi 等人^[7]提出一种时间序列不规则性和复杂度衡量指标散布熵 (dispersion entropy, DE),它具有计算高效的特点,并且考虑了时间序列中幅值间的相互关系。模糊熵 (fuzzy entropy, FE) 则在传统熵的基础上引入了模糊隶属度函数的概念,也能够准确地描述非线性特征^[8]。文献[9]提出了一种结合

经验模态分解并对近似熵进行了改进的辐射源识别方法,并运用支持向量机实现了辐射源的有效分类。

但在低信噪比情况下,通过经验模式分解会产生模式混淆、端点效应等问题。2014 年, Dragomiretskiy 等人^[10]提出了一种创新的信号处理算法变分模态分解 (variational mode decomposition, VMD),该算法具备自适应特性且完全非递归,能够有效解决模态混叠等问题。VMD 处理的模态分量具有不同的最优中心频率和有限的带宽,从而提高了信号的信噪比,使得不同模态分量间的特征差异更加显著,这有利于选取包含更多细微特征的信号分量。文献[11]利用 VMD 将接收到的信号同时分解为各种时间和频谱模式,证明了 VMD 在通信辐射源个体识别方面的鲁棒性。但其运用 VMD 分解信号时,没有对分解层数 K 这个关键参数做很好的解释,其大小关系到 VMD 算法分解的性能。文献[12]采用巴氏距离 (Bhattacharyya distance, BD) 被用作衡量两个不同概率分布之间相似度的一种有效方法,提出了选择最多样化特征的新方法,证明了巴氏距离在测量两个样本间相关程度的鲁棒性。文献[13]对信号进行 EMD、VMD 分解后的模态分量利用皮尔逊系数 (Pearson correlation coefficient, PCC) 的大小来衡量模态分量与原始数据之间的相关性,通过利用相关性分析 (correlation analysis, CA) 来剔除冗余的模态分量,进而提升了模型的预测准确性,这一过程凸显了进行相关性分析的必要性。

通信辐射源的个体识别特征,源于信号发射元件在传输数据时附加的微弱“噪声”信号。这些噪声会在信号传输中引入细微的高频成分,从而真正影响辐射源个体识别效果的是携带不同通信内容的低频数据信号^[14]。通过从辐射源发射的原始信号中去除非必要的成分,仅保留其独特的指纹特征,能够显著提升分类器的性能。基于以上讨论,以往研究运用 VMD 作为信号处理方法的时候,其分解层数是人为设定的,缺乏普适性和可靠性。并且没有进行细微特征分量的选取。因此本文在信号预处理部分利用巴氏距离,作为评估

各本征模态函数之间相似程度的方法, 无需人为设定分解层数, 来获得 VMD 算法的最优分解层数, 继而通过 VMD 对通信辐射源信号的每个符号波形信号进行分解得到各本征模态函数. 然后计算各本征模态函数与原始信号的相关系数, 去除数据信号, 选取包含更多细微指纹特征的高频模态分量. 接着在特征提取方面本文结合了时域、频域和熵特征参数, 并拼接融合成

多域特征向量实现了对辐射源的特征提取, 最后使用 LSTM 完成辐射源信号的分类识别.

1 通信辐射源发射机指纹建模

图 1 为构建的通信辐射源发射机通用模型, 并指明了射频指纹的产生来源. 下面将详细分析这些射频指纹产生的机理及其相应的建模策略.

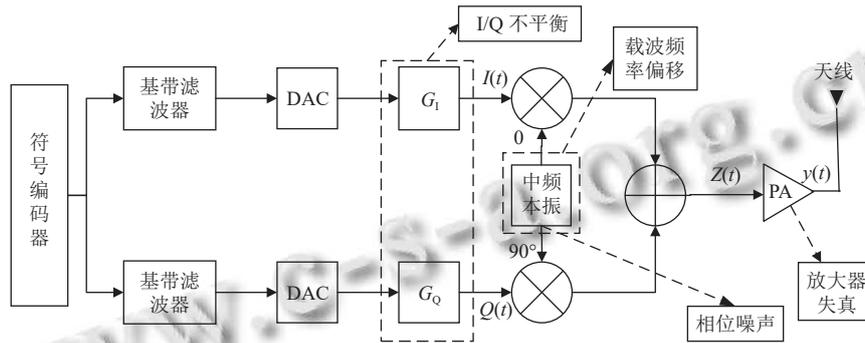


图 1 通信辐射源信号发射机模型

经过滤波器和数模转换处理后的发送符号, 首要步骤是进行 I/Q 正交调制. 在此过程中, 由于 I 路和 Q 路所使用的元器件存在个体差异, 经常会导致增益或相位上的不匹配, 从而引发 I/Q 不平衡现象. G_I 和 G_Q 分别代表 I/Q 两路的增益, 则两路的增益比为 $G_{I/Q} = G_I/G_Q$. 则 I/Q 两路产生的信号分别为:

$$I(t) = G_{I/Q} \sum_{k=-\infty}^{+\infty} I_k h(t - kT - \tau) \quad (1)$$

$$Q(t) = \sum_{k=-\infty}^{+\infty} Q_k h(t - kT - \tau - \tau_D) \quad (2)$$

其中, I_k 和 Q_k 分别代表的是符号编码器生成的两路编码, $h(t)$ 是基带滤波器的冲激响应, T 为符号周期, τ 是延迟失配, τ_D 表示 Q 路相对于 I 路的额外延迟. 之后经过正交调制得到中频信号:

$$Z(t) = I(t) \cos(\omega_c t + \zeta/2) + Q(t) \sin(\omega_c t - \zeta/2) \quad (3)$$

其中, ζ 表示相位不平衡参数, ω_c 为中频本振生成的频率.

由于不同发射机与接收机的晶体振荡器存在差异, 它们所产生的频率偏差绝对值也会各不相同. 同时中频本振还会产生相位噪声, 这都使得信号带有不同的指纹特征. 则中频信号的表示将变为:

$$Z(t) = I(t) \cos(\omega_c(1 + \varepsilon)t + \zeta/2 + \phi \cos \varpi t) + Q(t) \sin(\omega_c(1 + \varepsilon)t - \zeta/2 + \phi \cos \varpi t) \quad (4)$$

其中, ε 表示载频偏移程度, ϕ 和 ϖ 分别表示相位噪声的

噪声幅度和噪声频率.

另外, 在通信辐射源发射机中, 信号在经过功放时会产生一定的非线性^[15]. 在功率放大器的非线性程度较低的情况下, 采用 Taylor 级数模型进行建模能够更有效地描述和反映功放的非线性特征. 则发射机功放的输出信号可以表示为:

$$y(t) = a_0 + a_1 Z(t) + a_2 Z^2(t) + \dots = \sum_{m=0}^M a_m Z^m(t) \quad (5)$$

其中, a_m 为功放 Taylor 级数模型的第 m 阶参数, a_m 的不同直接导致了不同功率放大器的非线性差异.

功放中各谐波成分之间存在着较多的互调成分, 尤其是三阶互调成分对射频信号的影响最为显著, 假定功放输出的三阶分量表示为:

$$y_D^3 = a_3 (V_x \cos \omega_x t + V_c \cos \omega_c t)^3 \quad (6)$$

其中, V_x 表示信号的幅度, ω_x 为信号频率, V_c 表示被耦合的本振信号的振幅. 则接收端信号表示为:

$$r(t) = \sum_{m=0}^M a_m Z^m(t) + n(t) \quad (7)$$

其中, $n(t)$ 代表高斯白噪声.

根据式 (4) 可知, 由式 (6) 表示的三阶互调成分与图 1 中发射机细微特征的影响因素相关, 该三阶互调成分位于发射信号频带中, 与射频信号频率太相近, 经功率放大后的带通滤波很难将其滤掉, 从而形成一个

非线性细微指纹特征, 并以此作为个体特征来区分不同的辐射源. 所以, 在发射机的全通道中, 有许多具有时间域、频域和非线性特征的模块, 通过对这些特征的提取, 可以实现对单个辐射源的识别.

2 BDmin-VMD-CA 通信信号符号波形分解并选取细微特征分量算法

2.1 VMD 分解通信信号算法原理

通信辐射源信号具有非平稳和非线性特性, VMD 算法特别适合处理这种信号, 其分解过程就是将辐射源信号 $r(t)$ 的第 l 个符号周期的接收波形 $r_l(t)$, 分解成 K 个具有调幅和调频特性的窄带信号 $u_k^l(t)$, 即:

$$r(t) = \sum_{l=1}^L r_l(t) = \sum_{l=1}^L \sum_{k=1}^K u_k^l(t) = \sum_{l=1}^L \sum_{k=1}^K A_k^l(t) \cos(\varphi_k^l(t)) \quad (8)$$

其中, $A_k^l(t)$ 表示的是信号的瞬时幅值, $\varphi_k^l(t)$ 表示为相位函数.

给定通信信号符号波形 $r_l(t)$, 找到 K 个窄频带的 IMF 分量 $u_k^l(t)$, 且每个分量的中心频率为 ω_k^l , 对应的约束变分模型为:

$$\min_{\{u_k^l, \omega_k^l\}} \left\{ \sum_{k=1}^K \left\| \partial(t) \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k^l(t) \right] e^{-j\omega_k^l t} \right\|_2^2 \right\} \quad (9)$$

其中, $(\delta(t) + j/\pi t) \times u_k^l(t)$ 是对模态函数 $u_k^l(t)$ 进行 Hilbert 变换并求得的解析信号.

引入二次惩罚因子 α 和拉格朗日乘算符 $\lambda(t)$, 具有约束条件的变分法转化成一个不受限制的变分问题, 得拉格朗日方程式为:

$$L(\{u_k^l\}, \{\omega_k^l\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial(t) \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k^l(t) \right] e^{-j\omega_k^l t} \right\|_2^2 + \left\| r_l(t) - \sum_{k=1}^K u_k^l(t) \right\|_2^2 + \langle \lambda(t), r_l(t) - \sum_{k=1}^K u_k^l(t) \rangle \quad (10)$$

接着通过交替更新 $(u_k^l)^{n+1}$, $(\omega_k^l)^{n+1}$ 以及 λ^{n+1} , 寻求式 (10) 的鞍点. 当判别精度满足条件 $\varepsilon > 0$ 时, 循环过程终止.

$$\frac{\sum_{k=1}^K \left\| (U_k^l)^{n+1}(\omega) - (U_k^l)^n(\omega) \right\|_2^2}{\left\| (U_k^l)^{n+1}(\omega) \right\|_2^2} < \varepsilon \quad (11)$$

其中, $U_k^l(\omega)$ 为 $u_k^l(t)$ 的傅里叶变换, 本文取 $\varepsilon = u_k^l(t)$.

2.2 BDmin-VMD 通信信号符号波形分解算法

在统计分析领域, 巴氏距离被用作衡量两个离散或连续概率分布之间相似程度的一种手段, 可被用来测量两个样本间的相关程度. 当两个 IMF 成分间不包含同一时序频率, 也就是没有出现频率混叠现象, 其巴氏距离值较小; 反之, 其巴氏距离值较大, 且趋近于 0. 为优化 VMD 分解结果, 提出利用最小巴氏距离 (BDmin) 确定 VMD 分解层数 K 值.

该方法首先初始化 VMD 分解层数 $K=2$, 计算各相邻 IMF 分量之间的 BD 值. 而 $u_k(t)$ 并不满足归一化条件, 不是严格意义上的概率密度函数, 所以对各个相邻本征模态函数 $u_k^l(t)$ 与 $u_{k+1}^l(t)$ 取离散样本并进行归一化得到 $\tilde{u}_k^l[n]$ 与 $\tilde{u}_{k+1}^l[n]$, 表示为:

$$\tilde{u}_k^l[n] = \frac{u_k^l[n]}{\sum_{n=0}^{N-1} u_k^l[n]} \quad (12)$$

则巴氏系数 BC 表示为:

$$BC(\tilde{u}_k^l, \tilde{u}_{k+1}^l) = \sum_{n=0}^{N-1} \sqrt{\tilde{u}_k^l[n] \cdot \tilde{u}_{k+1}^l[n]} \quad (13)$$

最后计算各个相邻本征模态函数 $u_k^l(t)$ 与 $u_{k+1}^l(t)$ 之间的巴氏距离, 定义为:

$$BD(\tilde{u}_k^l, \tilde{u}_{k+1}^l) = -\log(BC(\tilde{u}_k^l, \tilde{u}_{k+1}^l)) \quad (14)$$

提取其中的 BDmin 值, 接着在每一次循环中将分解层数 K 值加 1, 更新 VMD 分解后提取的 BDmin 值, 在 BDmin 低于临界值的情况下 (通过大量的实验分析, 将该阈值设为 0.001, 趋于 0), 结束循环, 最终分解层数值取 K , 对通信辐射源符号波形进行分解.

2.3 基于 CA 的细微特征分量选取

VMD 算法对信号分解后得到多个 IMF 分量, 为了选出包含更多个体指纹特征的细微特征分量, 利用 PCC 的大小来衡量分解得到的各 $u_k^l(t)$ 与原始信号之间的相关性, 计算第 k 个本征模态函数 $u_k^l(t)$ 与通信信号符号波形 $r_l(t)$ 间的 PCC. 计算公式为:

$$\rho_{r_l(t), u_k^l} = \frac{\sum_{l=1}^L [r_l(t) - \bar{r}_l(t)] [u_k^l(t) - \bar{u}_k^l(t)]}{\sqrt{\sum_{l=1}^L [r_l(t) - \bar{r}_l(t)]^2 \sum_{l=1}^L [u_k^l(t) - \bar{u}_k^l(t)]^2}} \quad (15)$$

一般认为, IMF 分量中, 那些包含更多原始信号特征信息的部分, 与原信号的相关性更强, 这些往往是代表数据信号的低频成分; 相反, 那些所含信号数据信息

较少的 IMF 分量, 则被视为噪声模态分量, 即包含更多个体细微指纹特征的高频信号分量. 将与原始信号相关性最大的模态分量作为数据信号分量去除, 其他模态分量作为提取个体指纹特征的细微特征分量, 并将其表示为 $u_{kH}^l(t)$.

2.4 基于 BDmin-VMD-CA 信号分解并选取细微特征分量整体流程

本文的基于 BDmin-VMD-CA 分解信号并选取细微特征分量算法流程图如图 2 所示, 首先采用最小巴氏距离法, 通过计算不同 K 值下各本征模态函数间的巴氏距离值, 来获得 VMD 最优分解层数 K 对每一符号波形进行分解, 得到含有细微指纹特征的高频成分, 以及含有数据信息的低频成分; 然后计算各本征模态函数与其符号波形信号的相关系数, 选取相关系数小的本征模态函数作为辐射源细微特征分量, 完成基于 BDmin-VMD-CA 的信号预处理.

3 基于 LSTM 的多域细微特征提取与分类识别

3.1 时域细微特征

1) 包络细微特征: 反映了通信信号符号波形的脉冲包络振幅随时间的变化情况, 不同辐射源在调制方式和信号强度上可能存在差异, 包络特征能够有效捕捉这些差异, 可以作为辐射源个体识别的一个有效特征.

由于希尔伯特变换具有较好的适用性与抗噪性, 因此本文采用该方法提取包含细微特征的模态分量 $u_{kH}^l(t)$ 的包络, 将其与 $1/\pi t$ 卷积后进行 90° 移相, 表示为:

$$\hat{u}_{kH}^l(t) = u_{kH}^l(t) \times \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u_{kH}^l(\tau)}{t-\tau} d\tau \quad (16)$$

将移相后的信号 $\hat{u}_{kH}^l(t)$ 与原始信号 $u_{kH}^l(t)$ 构建成一个解析函数 $Z(t) = u_{kH}^l(t) + j\hat{u}_{kH}^l(t)$, 则 $u_{kH}^l(t)$ 的包络可表示为:

$$E_k(t) = \sqrt{|u_{kH}^l(t)|^2 + |\hat{u}_{kH}^l(t)|^2} \quad (17)$$

2) 瞬时频率和瞬时相位特征: 瞬时频率反映了通信信号符号波形频率随时间的变化情况, 并且具有一定的稳定性; 瞬时相位包含了通信信号符号波形的完整相位信息, 并且能够在一定程度上抑制噪声对通信辐射源识别的影响. 瞬时频率和瞬时相位特征可以揭示辐射源在频率调制和相位调制上的细微差异, 这些差异是辐射源指纹的重要组成部分.

瞬时相位是解析信号的复数角度, 即:

$$\theta_k(t) = \arctan \frac{\hat{u}_{kH}^l(t)}{u_{kH}^l(t)} \quad (18)$$

瞬时频率 $f_k(t)$ 表示信号在某一时刻的频率值, 通过对瞬时相位 $\theta_k(t)$ 求导来得到:

$$f_k(t) = \frac{1}{2\pi} \frac{d}{dt} \theta_k(t) \quad (19)$$

由式 (17)-(19) 可得时域融合特征向量 $T_1 = [E_k, \theta_k, f_k]$.

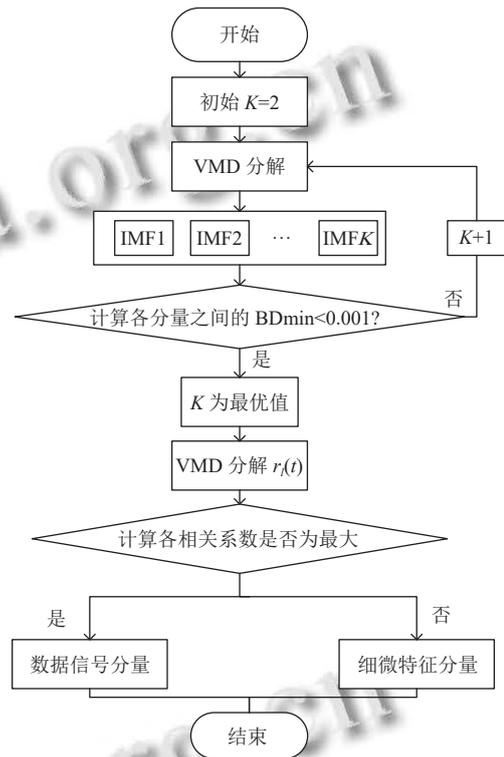


图 2 基于 BDmin-VMD-CA 信号分解并选取细微特征分量算法流程图

3.2 频域细微特征

1) FFT 频谱系数差分: 反映了通信信号符号波形在频域上的细微变化, 且增强了识别算法对噪声的鲁棒性, 使得在复杂通信环境中也能保持识别的稳定性和可靠性.

第 1 步, 对包含细微特征的模态分量 $u_{kH}^l(t)$ 进行采样, 得到离散信号 $u_{kH}^l(n)$. 第 2 步, 对 $u_{kH}^l(n)$ 进行 FFT 变换, 得到其在频域上的表示 $U_{kH}^l[n]$, 则 FFT 频谱系数差分表示为:

$$\Delta U_{kH}^l[n] = U_{kH}^l(n+1) - U_{kH}^l(n) \quad (20)$$

2) 功率谱上 3 dB 带宽内的频谱信息波形细节特征: 描述了通信信号符号波形在频域上的能量分布情况, 在 3 dB 带宽内, 信号的频谱信息相对稳定, 能够较好地反映辐射源的个体差异, 从而提高个体识别的准确性.

使用 Welch 周期图法来估计 $u_{kH}^l(t)$ 的功率谱密度, 该方法特别适用于处理非平稳信号和减少频谱泄露. 并针对信号 3 dB 带宽内的频域数据求统计特征: 均值、偏度和峰度. 假设用 $p(n)$ 表示功率谱数据, 则频谱的偏度 S 和峰度 F 定义为:

$$S_k = \frac{\frac{1}{N_p} \sum_{n=1}^{N_p} [|p(n)| - \bar{p}]^3}{\left\{ \frac{1}{N_p} \sum_{n=1}^{N_p} [|p(n)| - \bar{p}]^2 \right\}^{3/2}} \quad (21)$$

$$F_k = \frac{\frac{1}{N_p} \sum_{n=1}^{N_p} [|p(n)| - \bar{p}]^4}{\left\{ \frac{1}{N_p} \sum_{n=1}^{N_p} [|p(n)| - \bar{p}]^2 \right\}^2} \quad (22)$$

其中, N_p 是 3 dB 带宽内数据点的数量.

将这 3 种特征值组成频域融合特征向量 $T_2 = [\Delta U_k, S_k, F_k]$.

3.3 非线性细微特征

熵通常被视为描述数据或系统非线性特征的一种度量, 可作为表征辐射源信号的细微特征参数. 长度为 N 的细微特征分量信号 $u_{kH}^l(n), n=1, 2, 3, \dots, N$, 相空间重建后, 相空间矩阵如下:

$$U_{kH}^l = \begin{bmatrix} u_{kH}^l(1) & u_{kH}^l(1+d) & \dots & u_{kH}^l(1+(m-1)d) \\ \vdots & \vdots & \vdots & \vdots \\ u_{kH}^l(i) & u_{kH}^l(i+d) & \dots & u_{kH}^l(i+(m-1)d) \\ \vdots & \vdots & \vdots & \vdots \\ u_{kH}^l(M) & u_{kH}^l(M+d) & \dots & u_{kH}^l(N) \end{bmatrix} \quad (23)$$

其中, $U_{kH}^l(i) = \{u_{kH}^l(i), u_{kH}^l(i+d), \dots, u_{kH}^l(i+(m-1)d)\}, i=1, 2, \dots, N-(m-1)d, m$ 表示嵌入维数, d 表示时延, $M = N-(m-1)d$ 为相空间中点的个数.

1) 模糊熵特征

模糊熵是一种衡量信号复杂度和不规则性的指标, 能够反映信号的随机性和不可预测性. 模糊熵在熵的基础上引入了模糊隶属度函数, 可以对非线性特性进行更精确的描述. 其算法如下.

对于一个长度为 N 的时间序列 $u_{kH}^l(n)$ 进行相空间重建, 得到的时间顺序 Y_{kH}^l 表示为:

$$Y_{kH}^l(i) = \{u_{kH}^l(i), u_{kH}^l(i+d), \dots, u_{kH}^l(i+(m-1)d)\} - (u_{kH}^l(i))^0 \quad (24)$$

其中, $(u_{kH}^l(i))^0$ 为均值, 利用模糊函数计算 $Y_{kH}^l(i)$ 与 $Y_{kH}^l(j)$

之间的相似度:

$$D_{i,j}^m(q,r) = \mu(d_{i,j}^m, q, r) = \exp\left(-\frac{(d_{i,j}^m)^q}{r}\right) \quad (25)$$

其中, $d_{i,j}^m$ 为两个时间序列 $Y_{kH}^l(i)$ 与 $Y_{kH}^l(j)$ 之间的距离, $\mu(d_{i,j}^m, q, r)$ 为模糊函数, r 为相似容限度, q 决定了相似容限度边界的梯度. 定义函数:

$$\phi^m(q,r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{i,j}^m \right) \quad (26)$$

将维数变为 $m+1$, 重复之前步骤, 可得到 $\phi^{m+1}(q,r)$, 则其模糊熵可表示为:

$$FE_k(m, q, r, N) = \ln \phi^m(q, r) - \ln \phi^{m+1}(q, r) \quad (27)$$

在计算模糊熵时, 取 $m=2, q=2, r=0.2 \times std$, std 为 $u_{kH}^l(n)$ 的标准差.

2) 散布熵特征

散布熵是描述时间序列不规则性与复杂程度的一种新的度量方法, 它可以快速地反映出不同幅度间的相互关系.

对于细微特征分量序列 $u_{kH}^l(n), n=1, 2, 3, \dots, N$, 采用正态累积分布函数将 $u_{kH}^l(n)$ 映射到 $y_{kH}^l = \{y_{1kH}^l, y_{2kH}^l, \dots, y_{NkH}^l\}, y_{kH}^l \in (0, 1)$, 再通过线性变换将 y_{kH}^l 分配到 $[1, 2, \dots, C]$ 范围内的整数中, 即:

$$Z_j^C = \text{round}(C \cdot y_j + 0.5) \quad (28)$$

其中, Z_j^C 表示时间序列的第 j 个点, round 代表取整函数, 分类的数目用 C 表示. 然后创建嵌入向量 $Z_i^{m,C}$:

$$Z_i^{m,C} = (Z_i^C, Z_{i+d}^C, \dots, Z_{i+(m-1)d}^C) \quad (29)$$

每一时间序列 $Z_i^{m,C}$ 被映射到一个散布模式 $\pi_{v_0 v_1 \dots v_{m-1}}, v=1, 2, \dots, C$, 其中, $Z_i^C = v_0$. 并且有 C^m 种散布模式, 其相对频率 $p(\pi_{v_0 v_1 \dots v_{m-1}})$ 如下:

$$p(\pi_{v_0 v_1 \dots v_{m-1}}) = \frac{\text{Number}(\pi_{v_0 v_1 \dots v_{m-1}})}{N - (m-1)d} \quad (30)$$

其中, $\text{Number}(\pi_{v_0 v_1 \dots v_{m-1}})$ 是指嵌入向量 $Z_i^{m,C}$ 映射到 $\pi_{v_0 v_1 \dots v_{m-1}}$ 的数量.

根据 Shannon 熵的定义, 时间序列 $u_{kH}^l(n)$ 的散布熵可以定义为:

$$DE_k(m, C, d, N) = - \sum_{\pi=1}^{C^m} p(\pi_{v_0 v_1 \dots v_{m-1}}) \ln(p(\pi_{v_0 v_1 \dots v_{m-1}})) \quad (31)$$

设置散布熵的参数 $m=3, d=1, C=6$. 将这两个熵特征值拼接融合组成非线性特征向量 $T_3 = [FE_k, DE_k]$.

3.4 基于 LSTM 的分类识别

LSTM 作为分类识别模型,是专为处理序列数据而设计的。LSTM 能够充分利用提取的时域、频域和熵特征,捕捉到细微特征间的时序依赖关系,对辐射源信号每个符号的多域特征向量依次进行学习分类,从而实现对通信辐射源信号的准确分类。本文构建的 LSTM 由 1 个输入层,1 个 LSTM 层,1 个 dropout 层,1 个全连接层组成。LSTM 层具有 num_layers 层,每一层都包含 128 个 LSTM 单元。在 LSTM 层输出之后,应用了 1 个 dropout 层,其丢弃率设置为 0.3,以减轻过拟合的风险。最后全连接层将 LSTM 层的输出映射到最终的输出类别。网络训练次数为 50 次,Batchsize 设为 32,学习率设为 0.001 进行实验。基于 LSTM 的分类识别的整体流程图如图 3 所示,首先利用 BDmin-VMD 算法对通信辐射源符号波形进行预处理,得到最佳各信号分量,然后选取细微特征信号分量,提取时域、频域、模糊熵和散布熵特征参数,得到融合的时域特征向量 T_1 、频域特征向量 T_2 和熵特征向量 T_3 ,拼接融合成多域特征向量 $T = [T_1, T_2, T_3]$ 。其次将特征集分为训练集和测试集送入训练好的分类器中,得到分类识别结果完成通信辐射源个体识别。

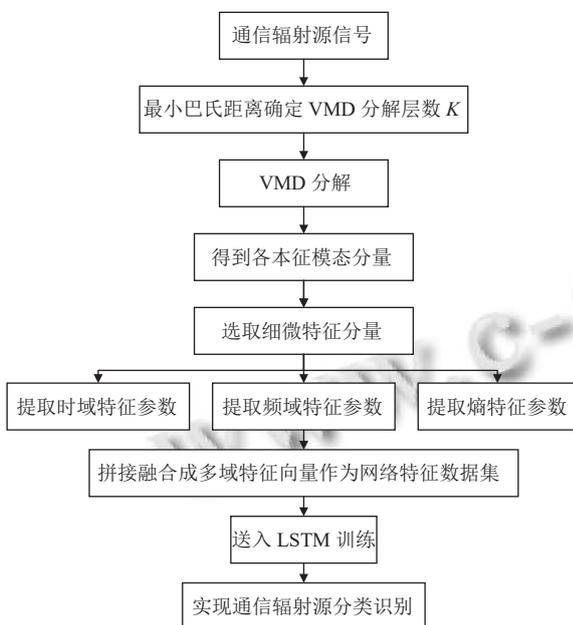


图 3 BDmin-VMD-CA 与 MFF 的通信辐射源个体识别方法流程图

4 实验仿真与结果分析

4.1 数据准备

实验所使用的数据集为开源数据集 Oracle,以

USRPX310 无线发射机为通信辐射源^[2],从数据集中选取 10 台 OFDM 发射机信号进行实验。每台 OFDM 发射机信号提取 1000 个样本,分别进行信号预处理,并将融合后的多域特征向量作为特征数据集。在此基础上,依据训练集和测试集按 8:2 的比例划分特征数据集,进行仿真模拟实验。信号预处理是基于 Matlab 2023a 平台实现的,神经网络的训练是基于 Python 3.9 和 PyTorch 1.9 实现的。考虑到神经网络学习过程中存在的随机因素,本文采用了 10 次重复实验并求平均值的方式,保证了实验结果的稳定可靠。

4.2 BDmin-VMD 算法分析

为验证本文提出利用 BDmin 自适应确定 VMD 分解层数 K 值方法所设定迭代终止条件成立,即当 BDmin 值小于 0.001 时,结束循环。以信噪比为 0 dB、2 dB、4 dB 和 6 dB 的 OFDM 通信辐射源信号为例,各选取其 100 个符号样本进行处理。

从图 4 中可以看出当 K 值取 5, 6, 7, 8 时, BDmin 值不在显著变化,这表明, VMD 分解过程中,两个 IMF 分量具有类似的成分,也就是过分解。由此证明了 BDmin-VMD 算法的迭代结束条件成立,可以将 VMD 分解的层数确定为 4,进而对原信号进行 VMD 分解。

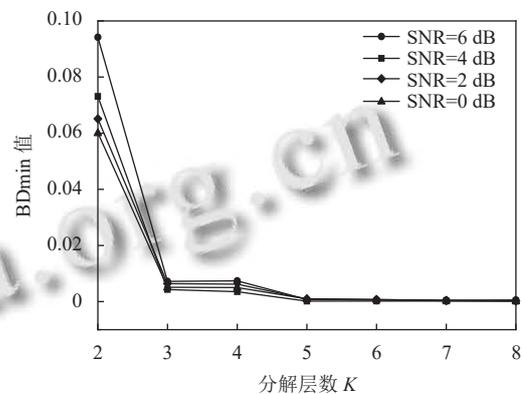


图 4 BDmin 值随 K 值变化图

从图 5 中可以看出将 OFDM 通信辐射源符号波形进行 BDmin-VMD 处理,连续将信号的低频和高频部分分开。图 5 中 IMF1-IMF4 是 OFDM 符号波形分解得到的第 1-4 个本征模态函数。

4.3 BDmin-VMD 算法验证

为验证本文所提 BDmin-VMD 方法的可靠性,对通信辐射源信号添加 SNR 取值区间为 [0, 6] dB,即能够覆盖从极低信噪比(接近噪声水平)到相对较好信噪比(能够清晰识别信号)的广泛范围,逐步增加 2 dB 的高斯白噪声对分解得到的各模态分量采用本文提出的

多域特征融合方法进行处理. 依据上述设定的参数, 对网络模型进行训练, 并评估算法在不同信噪比条件下

的表现. 从而达到验证 BDmin-VMD 算法分解信号效果是否可靠的目的, 识别结果如表 1.

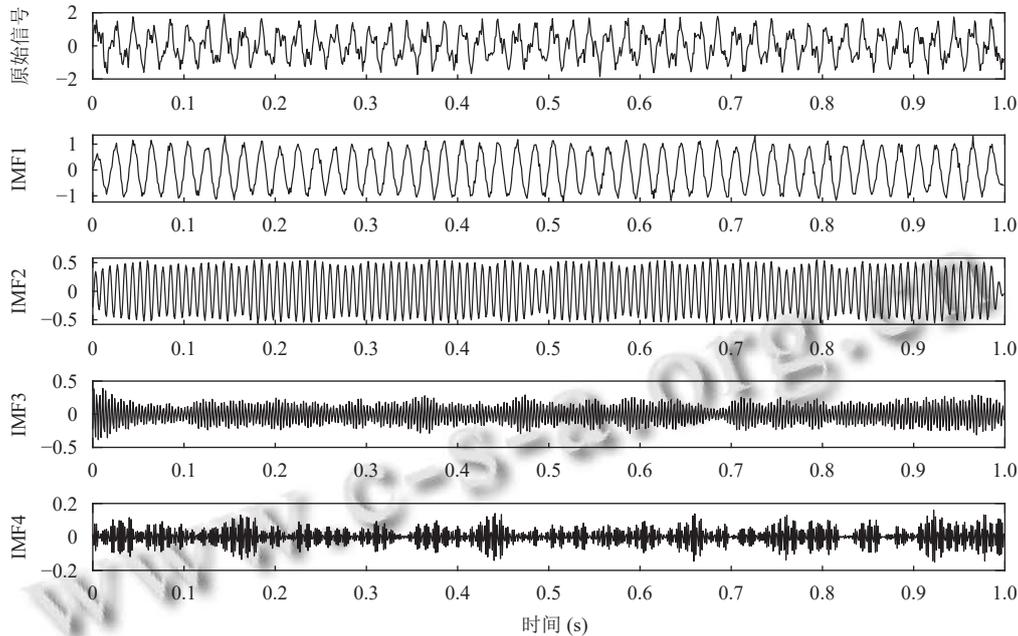


图 5 VMD 分解的各 IMF 图

表 1 不同信噪比下各分解层数的识别率 (%)

层数	0 dB	2 dB	4 dB	6 dB
2	44.5	65.6	72.6	76.1
3	51.2	70.5	81.0	82.9
4	72.6	82.9	90.1	92.7
5	58.3	79.1	86.4	90.2
6	53.8	73.8	84.5	89.3
7	40.3	62.2	74.9	80.6

由表 1 可知, VMD 分解性能受 K 值影响较大. 在不同信噪比的情况下, 当 $K=4$, 即 VMD 分解 4 层时, 通信辐射源个体识别率最高, 因此, 本文提出的 BDmin-VMD 信号分解算法是有效的.

4.4 验证基于相关性分析的模态分量选取对识别性能影响

由表 2 可知, 对通信辐射源符号波形进行 BDmin-VMD 处理后, 得到的第一本征模态函数与原始信号的相关系数最大, 认为 IMF1 分量为数据信号, 而 IMF2、IMF3 和 IMF4 分量与原始信号的相关系数数值皆远小于 IMF1, 将其作为包含更多细微特征的高频信号分量.

表 2 各 IMF 分量和原始信号之间的相关系数值

各 IMF 分量	IMF1	IMF2	IMF3	IMF4
相关系数值	0.7067	0.1294	0.0963	0.0034

在本实验中, 为验证相关性分析选取模态分量对识别性能的影响, 将 $K=4$ 时所有模态分量与去除 IMF1

分量的分解信号分别采用本文提出的多域特征融合方法进行处理后送入 LSTM 模型中分类识别, 其测试混淆矩阵如图 6 和图 7 所示. 从图中可以看出, 若将全部信号成分都进行特征提取后分类识别, 则会增大算法的复杂性, 降低识别率. 去除 IMF1 分量, 只保留高频细微特征分量进行特征提取并分类识别的识别准确率为 96.7%, 比对所有信号分量进行特征提取的识别准确率提高了 4%. 因此, 对各模态分量与原始信号进行相关性分析来选取细微特征分量可以提高通信辐射源信号的识别准确率.

4.5 验证各特征参数的有效性及对识别结果的影响

为了验证本文所提取的各特征参数在通信辐射源个体识别中的有效性, 本节设计了对比实验, 评估各特征参数对整体识别性能的影响. 对通信辐射源信号按本文所提方法分解并去除 IMF1 分量后添加 [0, 2, 4, 6] dB 的高斯白噪声, 依次去除某一特征参数与多特征参数融合进行对比实验, 所有实验均按照上述网络参数进行训练, 且不改变训练样本与测试样本的比率. 识别结果如表 3.

由表 3 可以看出, 去除某一特征参数后, 识别准确率均有所下降, 表明每个特征参数都对识别结果有一定的贡献, 验证了本文所选特征参数的有效性.

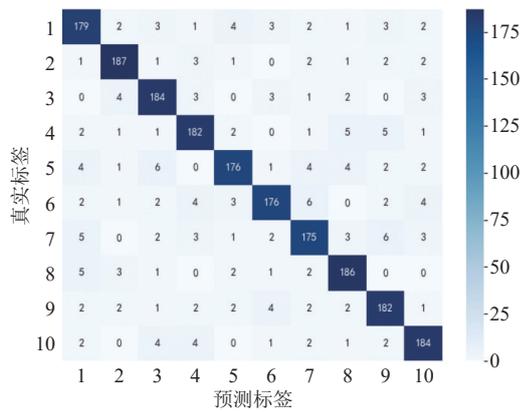


图6 所有分量提取特征分类识别测试混淆矩阵

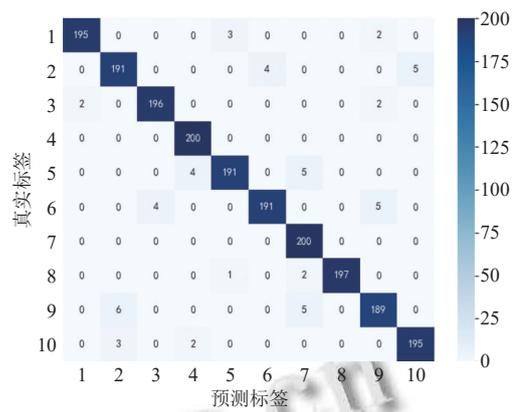


图7 去除 IMF1 分量后提取特征分类识别测试混淆矩阵

表3 多特征参数融合与单特征去除的识别准确率对比 (%)

参数	0 dB	2 dB	4 dB	6 dB	平均识别率
多特征参数融合	74.1	84.9	91.1	96.7	86.7
去除包络特征参数	68.3	79.2	87.5	90.1	81.3
去除瞬时频率和瞬时相位特征参数	62.4	72.8	80.3	84.6	75.0
去除FFT频谱系数差分特征参数	70.1	80.5	88.2	91.3	82.5
去除3 dB带宽内频谱特征参数	69.8	79.7	87.9	91.0	82.1
去除模糊熵特征层参数	67.2	71.6	79.3	85.0	75.8
去除散布熵特征参数	67.8	72.3	80.9	87.8	77.2

4.6 不同分类器对识别结果的影响

为进一步验证本文所提出的 LSTM 对辐射源分类识别的性能, 针对经过本文提出的信号预处理和特征提取后的 10 类通信辐射源样本数据, 选取了文献[16]中的支持向量机 (SVM)、文献[17]中的卷积神经网络 (CNN) 和随机森林 (RF) 分类算法进行对比实验, 以添加高斯白噪声的方式来改变信噪比, 进行多次重复实验, 根据每个分类器的分类识别结果统计出其平均识别率, 以减少误差对实验的影响. 不同信噪比下不同的分类器的识别率对比如图 8 所示.

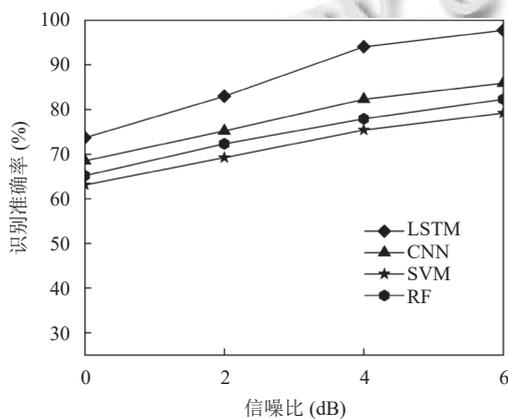


图8 不同信噪比下不同分类器的识别率对比曲线图

通过对比实验可以看出, 在信噪比为 6 dB 的条件

下, LSTM 作为识别分类器的识别准确率为 96.7%, 信噪比从 6 dB 降低到 0 dB 的过程中, LSTM 分类器在任何信噪比条件下, 识别效果都优于其他分类器, 可以说明在本文设计的通信辐射源个体识别实验中, LSTM 在通信辐射源个体识别任务中表现最优, 进一步证明了 LSTM 在处理时序特征方面的优势, 验证了本文选择 LSTM 作为分类器的合理性.

4.7 多特征参数融合分类识别结果及对比分析

为了验证本文方法的识别性能, 进行了 3 个对比实验. 所有方法均在信噪比为 6 dB, 且在不改变训练样本与测试样本比率的情况下, 按照上述参数进行训练.

在第 1 个对比实验中, 将本文方法 (即 BDmin-VMD-CA+多域特征) 与① BDmin-VMD-CA+时域特征、② BDmin-VMD-CA+频域特征、③ BDmin-VMD-CA+非线性特征 (熵特征) 进行多次重复比较实验, 并根据网络分类的结果统计了识别准确率随网络训练轮次的变化曲线. 得到的对比变化曲线如图 9 所示.

在第 2 个对比实验中, 将本文方法与① BDmin-VMD-CA+时域特征与频域特征、② BDmin-VMD-CA+时域特征与非线性特征 (熵特征)、③ BDmin-VMD-CA+频域特征与非线性特征 (熵特征) 同样进行多次重复比较实验, 得到的识别准确率随网络训练轮次的变化曲线如图 10 所示.

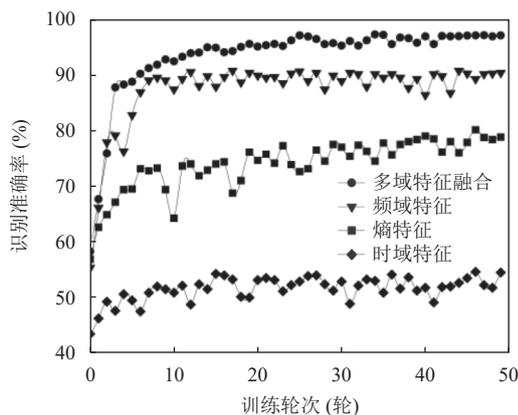


图9 单一变换域和多域特征融合的识别结果比较

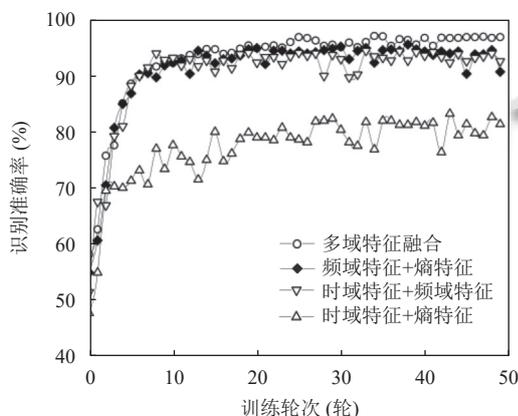


图10 基于两种变换域和多域特征的融合识别结果比较

通过比较实验表明,在10个OFDM仿真信源中,所采用的多域特征融合算法具有最好的识别精度.在网络训练20轮之后逐步趋于收敛,收敛后的平均识别准确率为96.7%.图9显示,仅依赖时域特征或熵特征时,训练结果不稳定且识别率不高,收敛后的平均识别准确率分别为55.3%和78.5%.因为时域特征虽能捕捉时间变化,但无法全面反映频率和能量分布;而熵特征信息也不足以区分信号.同样,频域特征虽强调频率信息,却忽视了时域和非线性特性,收敛后的平均识别准确率为90.1%.通过图10也可直观看出多域融合是提高个体识别性能的有效途径.并且只有将时域、频域和非线性域的特征参数进行融合,才可以综合考虑信号在不同方面的特性,提高识别性能和效果.

在第3个对比实验中,将本文方法与其他方法性能对比.将信号无分解+本文多域特征作为对比方法1.文献[18]对信号前导码的稳态部分提取了多维近似熵作为细微指纹特征,并对其进行分类.作为对比方法2.文献[16]首先用固有时间分解对辐射源信号进行分解,

然后提取了多个熵构成了特征向量,最后使用支持向量机技术完成辐射源分类识别,称为对比方法3.文献[17]采用固定边界的经验小波分解算法对辐射源信号进行处理后,将得到的子信号输入到卷积神经网络中进行分类,作为对比方法4.在保证训练样本与测试样本比率相同的情况下,将上述方法与本论文的方法进行比较,图11中的各识别结果是由各个方法中的分类器来得出的.

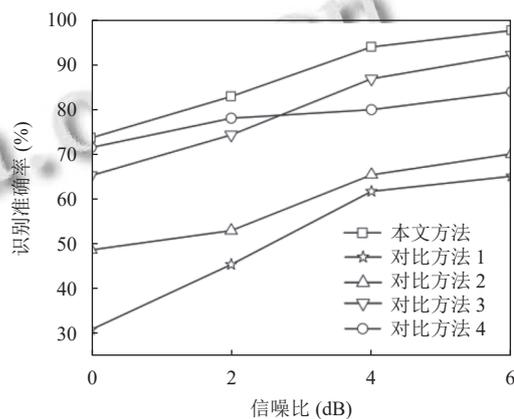


图11 各方法的识别率对比曲线

根据图11所示的信息,当信噪比为6 dB时,方法1对10台OFDM发射机的识别准确率为65.1%,相比之下,方法2至方法4针对同样数量的发射机的识别准确率分别为:71.8%、92.3%和84.2%.而本文所提算法识别准确度为96.7%.当信噪比下降到0 dB时,在本文提出的方法下,10台发射机的识别准确率达到74.1%,而与之对比的方法1-4对这些发射机的识别准确率则分别为:31.8%、49.1%、65.3%和71.2%.可以说明对通信辐射源符号波形进行本文提出的联合巴氏距离和相关性分析的变分模态分解预处理方法进行预处理后,再提取多域特征参数融合,识别准确率更高.

5 结论

本文提出了一种BDmin-VMD-CA+MDFF的通信辐射源个体识别方法.在信号预处理阶段提出了基于巴氏距离的VMD分解衡量得到的各个模态组分间的相似度.该方法无须人为设定VMD分解层数,可获得VMD分解的最优分解层数,确保通信辐射源符号波形信号完全分解.并结合相关性分析去除了包含更多数据信息的低频信号分量,只留下包含更多指纹特征的细微特征信号分量进行特征提取.特征提取阶段考虑

到单一域特征识别精度不高,对细微特征信号分量提取了多域特征参数并拼接融合成多域特征向量,最后使用长短期记忆网络完成分类识别.在公开数据集上的实验结果表明,本文提出的BDmin-VMD-CA+多域特征融合方法在信噪比为6 dB的条件下,对10台OFDM辐射源的识别准确率为96.7%,比各单一域平均识别准确率提高了22.1%,比不经过本文算法处理+多域特征融合方法识别准确率提高了31.6%.此外,本文使用数据集中的信号产生和传输环境均是在较为理想的实验室中进行的,然而在实际的应用场景中,接收到的通信辐射源信号会受到多种环境因素的影响,未来的工作可能会将该方法扩展到更复杂的信号环境中.

参考文献

- 1 Su JL, Liu H, Yang L. Specific emitter identification based on CNN via variational mode decomposition and bimodal feature fusion. Proceedings of the 3rd IEEE International Conference on Power, Electronics and Computer Applications (ICPECA). Shenyang: IEEE, 2023. 539–543.
- 2 刘高辉, 宋博武. DRSN与集成融合的OFDM辐射源个体识别方法. 信号处理, 2024, 40(6): 1062–1073.
- 3 Qu LZ, Liu H, Huang KJ, *et al.* Specific emitter identification based on multi-domain feature fusion and integrated learning. Symmetry, 2021, 13(8): 1481. [doi: 10.3390/sym13081481]
- 4 Zhang XL, Li TY. Specific emitter identification based on feature diagram superposition. Proceedings of the 7th International Conference on Integrated Circuits and Microsystems (ICICM). Xi'an: IEEE, 2022. 703–707.
- 5 Ying WW, Deng PF, Hong SH. Channel attention mechanism-based multi-feature fusion network for specific emitter identification. Proceedings of the 4th IEEE International Conference on Civil Aviation Safety and Information Technology (ICCASIT). Dali: IEEE, 2022. 1325–1328.
- 6 Huang GQ, Yuan YJ, Wang X, *et al.* Specific emitter identification based on nonlinear dynamical characteristics. Canadian Journal of Electrical and Computer Engineering, 2016, 39(1): 34–41. [doi: 10.1109/CJECE.2015.2496143]
- 7 Rostaghi M, Azami H. Dispersion entropy: A measure for time-series analysis. IEEE Signal Processing Letters, 2016, 23(5): 610–614. [doi: 10.1109/LSP.2016.2542881]
- 8 Chen WT, Wang ZZ, Xie HB, *et al.* Characterization of surface EMG signal based on fuzzy entropy. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2007, 15(2): 266–272. [doi: 10.1109/TNSRE.2007.897025]
- 9 Xie Y, Wang SL, Zhang EY, *et al.* Specific emitter identification based on nonlinear complexity of signal. Proceedings of the 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Hong Kong: IEEE, 2016. 1–6.
- 10 Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Transactions on Signal Processing, 2014, 62(3): 531–544. [doi: 10.1109/TSP.2013.2288675]
- 11 Satija U, Trivedi N, Biswal G, *et al.* Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios. IEEE Transactions on Information Forensics and Security, 2019, 14(3): 581–591. [doi: 10.1109/TIFS.2018.2855665]
- 12 Shah MH, Dang XY. Novel feature selection method using Bhattacharyya distance for neural networks based automatic modulation classification. IEEE Signal Processing Letters, 2020, 27: 106–110. [doi: 10.1109/LSP.2019.2957924]
- 13 刘梦, 韩建, 曹志民, 等. 基于模态分解与LSTM注意表征的测井曲线重构研究. 电子学报, 2024, 52(4): 1399–1410. [doi: 10.12263/DZXB.20220887]
- 14 徐晋凯, 谢钧, 俞璐. 基于VMD的通信辐射源个体识别方法. 通信技术, 2022, 55(4): 461–467. [doi: 10.3969/j.issn.1002-0802.2022.04.008]
- 15 Zhang JQ, Woods R, Sandell M, *et al.* Radio frequency fingerprint identification for narrowband systems, modelling and classification. IEEE Transactions on Information Forensics and Security, 2021, 16: 3974–3987. [doi: 10.1109/TIFS.2021.3088008]
- 16 任东方, 张涛, 韩洁. 结合ITD与非线性分析的通信辐射源个体识别方法. 信号处理, 2018, 34(3): 331–339.
- 17 Bremnes K, Moen R, Yeduri SR, *et al.* Classification of UAVs utilizing fixed boundary empirical wavelet sub-bands of RF fingerprints and deep convolutional neural network. IEEE Sensors Journal, 2022, 22(21): 21248–21256. [doi: 10.1109/JSEN.2022.3208518]
- 18 Sun LT, Wang X, Yang AF, *et al.* Radio frequency fingerprint extraction based on multi-dimension approximate entropy. IEEE Signal Processing Letters, 2020, 27: 471–475. [doi: 10.1109/LSP.2020.2978333]

(校对责编: 张重毅)