



# 基于多域判别核典型相关分析的辐射源指纹特征融合方法

孙丽婷, 王翔\*, 黄知涛

国防科技大学电子科学学院, 长沙 410073

\* 通信作者. E-mail: christopherwx@163.com

收稿日期: 2021-06-04; 修回日期: 2021-08-21; 接受日期: 2021-12-20; 网络出版日期: 2023-01-06

湖南省创新群体研究项目 (批准号: 2019JJ10004) 和国防科技大学青年创新奖 (批准号: 18/19-QNCXJ) 资助项目

**摘要** 辐射源个体识别 (specific emitter identification, SEI) 是指通过提取信号中蕴含的有关其发射来源的硬件指纹信息, 来实现对特定信号辐射源的精确识别. SEI 技术的关键在于指纹特征的提取. 相关研究大多侧重于定义和提取新的指纹特征, 较少关注对已有特征的综合利用问题. 鉴于不同分析域的特征对辐射源指纹的描述存在互补性, 本文提出一种基于多域判别核典型相关分析 (multi-domain discriminant kernel canonical correlation analysis, MDKCCA) 的辐射源指纹多域特征融合方法, 充分利用特征的标签信息以及特征间的互补性, 在高维空间完成多域特征的降维与融合. 以 4 个特征分析域 8 种常见指纹特征为依托, 在 4 种不同类型的实测数据集上验证了算法的性能. 结果证明, 该方法无需人工特征寻优环节, 可大幅降低融合特征的维度, 对 4 类目标的准确识别率均达到 95% 以上, 优于最优单一特征, 同时优于基于直接级联或基于 PCA (principal component analysis) 降维变换的简单特征综合方法、基于神经网络的特征综合方法, 以及基于判别相关分析 (discriminant canonical correlation, DCA) 等方法的特征融合方法.

**关键词** 辐射源个体识别, 特征融合, 多域辐射源指纹特征, 典型相关分析, 特征提取

## 1 引言

辐射源个体识别技术 (specific emitter identification, SEI) 通过测量信号的外部特征, 提取信号中蕴含的发射机硬件信息, 对发射信号的来源进行精确识别<sup>[1,2]</sup>. 这种硬件信息是指由于发射机在生产以及工作过程中存在各种非理想特性, 导致发射的信号不是严格意义上的完美信号, 不可避免地携带硬件的非理想特性<sup>[3]</sup>. 这种特性因设备而异, 如同人的指纹, 具有明显的个体色彩, 相对稳定, 且无法

**引用格式:** 孙丽婷, 王翔, 黄知涛. 基于多域判别核典型相关分析的辐射源指纹特征融合方法. 中国科学: 信息科学, 2023, 53: 146–163, doi: 10.1360/SSI-2021-0197  
Sun L T, Wang X, Huang Z T. Radio frequency fingerprinting feature fusion based on multi-domain discriminant kernel canonical correlation analysis (in Chinese). Sci Sin Inform, 2023, 53: 146–163, doi: 10.1360/SSI-2021-0197

消除<sup>[4]</sup>. 因此, 该技术又被称为射频指纹识别 (radio frequency fingerprinting, RFF)<sup>[5]</sup>, 在频谱管理、认知无线电<sup>[6]</sup>、无线网络安全<sup>[7]</sup> 等诸多领域都有广泛的应用.

SEI 的基本处理流程包括数据接收、数据预处理、指纹特征提取、特征综合, 以及分类识别. 而其中核心的步骤是指纹特征提取<sup>[8]</sup>. 迄今为止研究人员已经提出诸多指纹特征, 可以分为时域特征<sup>[9~11]</sup>、频域特征<sup>[12]</sup>、时频域特征<sup>[13,14]</sup>、双谱域特征<sup>[15~17]</sup> 等. 近年来随着深度学习的火热, 基于神经网络自动提取指纹特征方法也越来越常见<sup>[18,19]</sup>.

现阶段指纹特征提取的研究主要关注单一特征, 或者是单一域特征. 其基本思路是通过在特定数据集上定义某一个最优特征以达到不错的识别结果. 然而, 不同特征对于不同信号、不同场景的表征能力不同, 适用范围和侧重点存在差异<sup>[20]</sup>. 仅用单一特征完成识别的思路缺乏通用性和泛化性. 实际上, 不同特征, 特别是不同域特征, 从不同角度提供了关于信号指纹的互不相关的新信息 (以下简称新息), 具备一定互补性<sup>[21]</sup>. 因此选择单一特征不是最优的思路, 综合利用效果非最优的其他特征可以进一步提升识别效果.

关于 SEI 特征综合利用的研究相对较少, 主要分为特征级融合以及决策级融合. 在决策级融合方面, 主要方法有随机森林<sup>[22]</sup>、多分类器决策等. 这些方法对分类器的选择和训练要求较高, 对不同特征之间的关系并未深入研究. 在特征级融合方面, 主要是将不同特征以串联形式直接输入到传统分类器<sup>[21]</sup>, 或者基于深度集成学习方法输入到神经网络<sup>[23]</sup> 中. 直接串联的特征维度往往较高, 且由于指纹特征差异比较细微, 导致高维空间内分布稀疏, “维度灾难” 和 “集中现象” 更加突出<sup>[24]</sup>. 此外在 SEI 实际应用中, 通常难以保证获取足够多的高质量标注数据, 不利于数据驱动的深度学习方法提高性能. 文献<sup>[25]</sup> 基于核典型相关分析 (kernel canonical correlation, KCCA) 实现了辐射源模糊域特征融合. 然而该方法进行的是单一域内不同特征切片的融合, 新息有限, 且只是从特征数据分布上进行压缩变换, 并未结合数据类别标注信息, 因此融合特征对识别性能改善有限. 特征综合利用应从更多的角度刻画辐射源的个体指纹, 充分发挥不同特征对辐射源指纹不同角度刻画能力的优势, 为 SEI 提供更加全面的信息<sup>[20]</sup>.

为解决上述问题, 本文对多个不同域的指纹特征进行综合处理, 提出一种基于多域判别核典型相关分析 (multi-domain discriminant kernel canonical correlation, MDKCCA) 的指纹特征融合方法. 通过核变换将原始特征进行非线性扩展, 同时结合辐射源个体标签信息进行类别约束下的特征优化融合. 本文的主要贡献如下:

第一, 证明了辐射源指纹特征之间的互补性. 同时, 充分利用新息并基于非线性映射变换寻找特征的非线性相关关系, 实现了对多种分析域特征的融合.

第二, 结合了数据的标注信息, 即基于特征类间散布特性构造特征优化的变换矩阵进行特征的融合, 融合后的特征对指纹的表征更加精确, 类间差异性更大.

第三, 在计算过程中, 直接将多域特征进行有效降维与综合利用, 省略了以往辐射源指纹识别应用中繁复的人工特征寻优环节, 能够适应不同的信号类型. 同时算法所需的超参数少, 易于设定, 大幅减少专家参与, 降低了应用门槛.

第四, 基于不同辐射源类型不同信号类型的 4 种实测数据集, 选择时域、时频域、模糊域、双谱域的 8 种经典指纹特征进行了实验. 结果表明该算法能够实现多域特征的有效综合, 在 4 种测试数据集上均能达到 95% 以上的正确识别率, 效果上优于最优单一特征识别方法、基于直接级联或 PCA (principal component analysis) 降维变换的简单特征综合方法、基于神经网络的特征综合方法, 以及基于判别相关分析 (discriminant canonical correlation, DCA) 等方法的特征融合方法, 提升了辐射源个体识别系统的性能.

本文其余章节结构如下: 第 2 节阐述 4 个指纹特征分析域 8 种特征的计算方法; 第 3 节通过实测数据, 实验验证分析特征间的互补性; 第 4 节给出本文算法原理以及其实现过程; 第 5 节在 4 种不同类型的实测数据集上实验, 验证本文所提多域特征融合方法的有效性; 第 6 节总结全文.

## 2 多域辐射源指纹特征

### 2.1 时域指纹特征

本文选用归一化信号包络以及信号的瞬时幅度作为时域 (T) 代表特征, 将平滑后的信号幅度作为瞬时幅度特征, 将归一化包络作为信号包络特征<sup>[9~11]</sup>, 分别用  $IA(t)$ ,  $Env(t)$  表示.

令  $x(t)$  表示待处理的信号, 若  $x(t)$  为实信号, 首先进行 Hilbert 变换  $\mathcal{H}[\cdot]$  得到复信号  $\hat{x}(t)$ ,

$$\hat{x}(t) = \mathcal{H}[x(t)] = \int_{-\infty}^{+\infty} \frac{x(\tau)}{\pi(t-\tau)} d\tau = x(t) * \frac{1}{\pi t}, \quad (1)$$

信号  $x(t)$  的包络可以依据下式计算:

$$a(t) = \sqrt{x^2(t) + \mathcal{H}^2[x(t)]}. \quad (2)$$

信号的平均功率为  $P = E[|a(t)|^2]$ , 其中,  $E[\cdot]$  表示期望计算. 因此, 信号归一化包络记为  $Env(t) = a(t)/P$ . 对  $a(t)$  进行平滑处理, 即得到瞬时幅度特征即  $IA(t) = \frac{1}{\omega_n} \sum_{i=0}^{\omega_n} |a(t+i)|$ , 其中,  $\omega_n$  为滑动窗宽.

### 2.2 时频域指纹特征

本文选择瞬时相位和瞬时频率作为时频域 (F) 指纹特征, 分别用  $IP$ ,  $IF$  表示. 瞬时相位和瞬时频率反映了辐射源系统的非线性特性, 常用于表征辐射源设备无意调制特性<sup>[26, 27]</sup>.

接收到的信号  $\hat{x}(t)$  的  $IP$  计算方法如下:

$$\varphi_0(t) = \text{angle}(\hat{x}(t)), \quad (3)$$

$$\tilde{\varphi}(t) = \frac{1}{2L+1} \sum_{l=-L}^L \varphi_0(t+l), \quad (4)$$

其中,  $2L+1$  是滑动平均窗的窗宽.

$$IP(t) = \varphi(t) = \frac{1}{\pi} (\tilde{\varphi}(t) - E[\tilde{\varphi}(t)]), \quad (5)$$

$IF$  为  $IP$  的导数:  $IF(t) = \frac{1}{2\pi} d\varphi(t)/dt$ .

### 2.3 模糊域指纹特征

模糊函数主要用于分析发射信号匹配滤波后的分辨性能, 也可用于表征信号的个体差异<sup>[28, 29]</sup>. 其定义为瞬时相关函数关于时间  $t$  的 Fourier 反变换<sup>[30]</sup>, 即

$$\mathbf{A}_x(\tau, \xi) = \int_{-\infty}^{\infty} \hat{x}(t)\hat{x}^*(t+\tau)e^{j2\pi\xi t} dt, \quad (6)$$

其中,  $(\cdot)^*$  表示信号的复共轭.

为去除模糊函数的冗余性, 通常将模糊函数频偏接近零 (包括零) 的切片作为指纹特征<sup>[25]</sup>. 本文选择模糊函数零频切片与第 1 切片作为指纹特征, 分别用  $Slice1$  和  $Slice2$  表示. 为保持表述的一致性, 文中模糊函数切片特征被称为模糊域 (S) 特征.

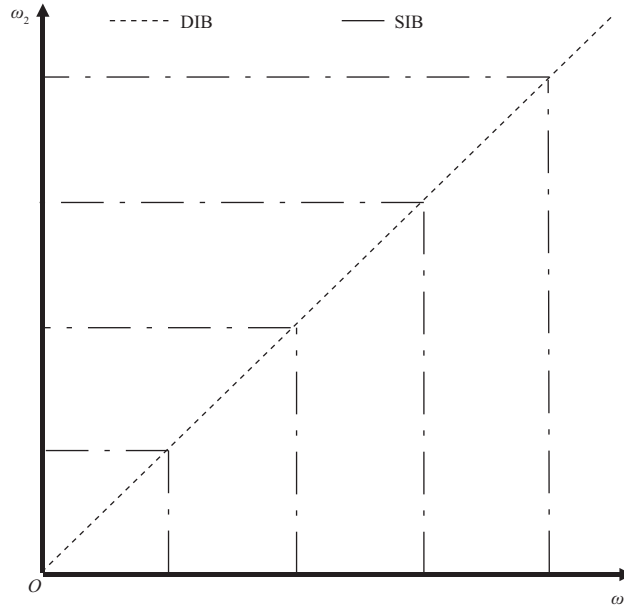


图 1 SIB 与 DIB 积分路径示意  
Figure 1 Integration paths of SIB and DIB

## 2.4 双谱域指纹特征

双谱是一种常见的高阶统计量, 具备时移不变性、相位保持性、尺度变化性, 理论上可以抑制高斯白噪声, 因此被广泛应用于辐射源个体识别问题<sup>[15,16]</sup>. 信号  $\hat{x}(t)$  的双谱为<sup>[31]</sup>

$$\mathbf{B}(\omega_1, \omega_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{C}_{3x}(\tau_1, \tau_2) \cdot \exp(-j(\omega_1\tau_1 + \omega_2\tau_2)) d\tau_1 d\tau_2, \quad (7)$$

其中,  $\mathbf{C}_{3x}(\tau_1, \tau_2) = E\{\hat{x}^*(t)\hat{x}(t+\tau_1)\hat{x}(t+\tau_2)\}$  为  $\hat{x}(t)$  的三阶相关系数,  $(\cdot)^*$  表示求信号的复共轭.

由于双谱维度过高, 在实际 SEI 应用中通常按照不同的路径对双谱  $\mathbf{B}(\omega_1, \omega_2)$  进行局部围线积分, 可以得到不同的双谱积分切片特征, 又叫做局部围线积分双谱<sup>[31]</sup>.

$$\mathbf{H}(\omega) = \oint_{S_l} \mathbf{B}(\omega_1, \omega_2) d\omega_1 d\omega_2, \quad (8)$$

其中,  $S_l$  表示不同的积分路径. 本文选择矩形积分双谱 (square integral bispectrum, SIB) 和对角切片双谱 (diagonal integral bispectrum, DIB) 作为双谱域 (B) 特征<sup>[20]</sup>, 分别用 SIB, DIB 表示. 对应的积分路径如图 1 所示.

## 3 多域指纹特征互补性分析

本节以 5 个相同型号的无线网卡发射的 IEEE802.11b 信号为例, 详细分析不同域指纹特征的互补特性. 5 个辐射源分别用 R1~R5 表示, 各选择 5 个测试样本, 共计 25 个. 分别对原始特征、单域多个特征决策融合、多域特征决策融合方法进行了实验分析.

(1) 原始特征方法分析. 从上文所述 4 种特征域中各选择一种特征作为原始特征, 分别为 Slice1, Env, IF 和 DIB, 识别结果如图 2 所示. 图中横坐标表示测试的样本序号, 每 5 个为一组, 来自同一

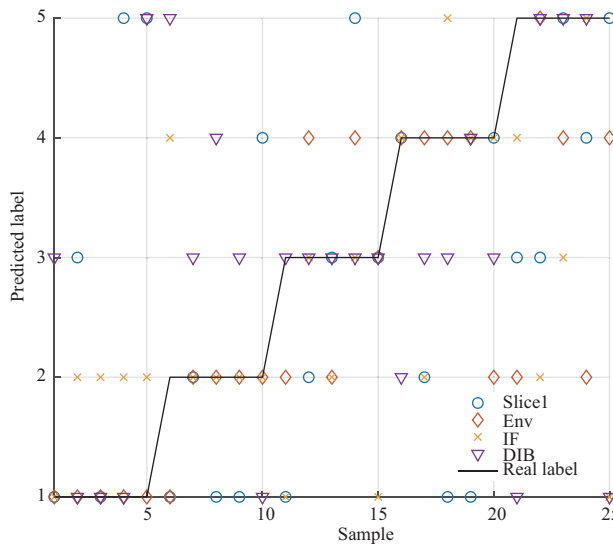


图 2 (网络版彩图) 原始特征方法识别结果

Figure 2 (Color online) Identification results of different features

个辐射源个体; 纵坐标表示特征方法预测的样本标签. 图中黑色曲线表示正确的标签信息, 不同形状颜色的点表示不同特征方法. 与黑线分离的散点表示错判样本, 结果与黑色曲线拟合的表示判别正确. 由图 2 可知, 不同特征方法错判点几乎不会重叠在一起, 即不同的特征呈现不同的错判规律, 表明不同特征对辐射源的指纹具有不同的观测视角.

**(2) 多域特征决策融合分析.** 综合考虑图 2 中 4 个域的特征, 采用基于投票的决策级融合方法, 即遵循少数服从多数的原则, 当票数相同时采用排序在前特征结果, 识别结果如图 3 所示. 由图 3(a) 可知, 与原始特征相比 (图 2), 多特征域的综合利用使得散布在黑线之外的错判样本数目大幅减少. 而对比不同辐射源的错误样本数 (图 3(b)) 可知, 相比于单一特征, 多域特征投票综合后平均错判样本数明显下降, 识别性能更加稳健. 因此, 证明了不同域的特征具备互补性, 综合利用可改善识别效果. 值得注意的是, 尽管多域特征投票性能总体上更加稳健, 然而由于未考虑特征自身的性能优劣, 所有特征以等比重的方式参与决策, 因此图 3(b) 中的投票结果对目标个体的而言并未优于最优特征.

**(3) 单域特征决策融合分析.** 将同一种域内的两种特征进行单域特征投票综合, 结果如图 4 所示. 与图 2 的单个特征识别效果相比, 同种域内的两个特征融合后错判样本减少, 识别结果有一定的提升, 但是明显低于图 3 的 4 种域特征综合结果. 例如, 图 4(c) 中模糊域特征融合后, 样本编号 10~20 (目标 R3~R4) 的样本全部判别错误, 而图 3 中这些样本绝大部分能够正确识别.

上述实验可以得到如下结论:

第一, 不同指纹特征对目标具有不同的观测视角.

第二, 指纹特征综合优于单一特征.

第三, 不同域指纹特征融合比单域特征融合提供更多新息, 识别效果更好.

第四, 不同的特征对识别的贡献度不同, 因此融合过程中应当考虑特征的优劣, 而不是将所有特征以等权重的方式进行融合.

因此, 本文基于多种域特征间的互补性并结合训练数据的真实类别标签, 完成多域特征综合利用. 在后文中, 将结合更多类型的数据对辐射源指纹特征互补性进行深入实验分析.

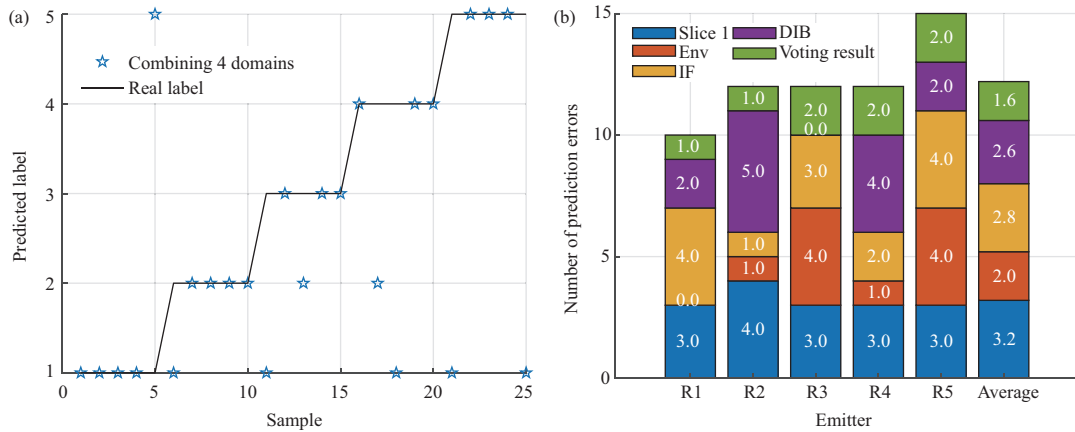


图 3 多域特征投票综合结果

Figure 3 Identification results of multi-domain features by voting. (a) Class predicted results; (b) distribution of error samples

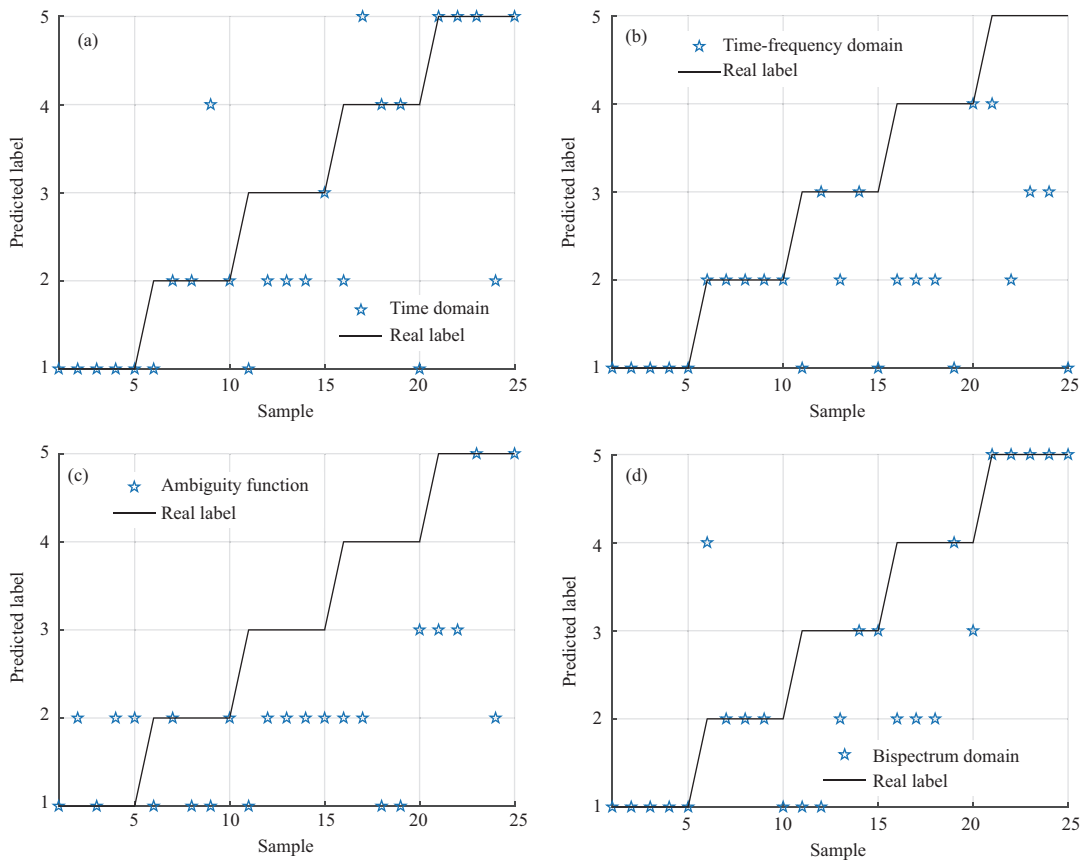


图 4 同域特征投票综合结果

Figure 4 Identification results of single-domain features by voting. (a) Domain T; (b) domain F; (c) domain S; (d) domain B

## 4 多域特征融合方法

### 4.1 典型相关分析

典型相关分析 (canonical correlation analysis, CCA), 又叫做典型关联分析, 能够从总体上揭示两组数据之间的内在关系, 常用于特征降维与特征融合 [32].

给定两个特征向量  $\mathbf{x} \in \mathbb{R}^p$ ,  $\mathbf{y} \in \mathbb{R}^q$ , 各有  $n$  个样本, 分别记作  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$  和  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n] \in \mathbb{R}^{q \times n}$ . 令  $\mathbf{w}_x, \mathbf{w}_y$  分别表示两个特征向量的投影方向, CCA 的目标是最大化式 (9) 的准则函数:

$$J(\mathbf{w}_x, \mathbf{w}_y) = \frac{\mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y}{\sqrt{\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y}}, \quad (9)$$

其中,  $\mathbf{C}_{xx} = \mathbf{X}\mathbf{X}^T$ ,  $\mathbf{C}_{yy} = \mathbf{Y}\mathbf{Y}^T$  分别为  $\mathbf{X}$  和  $\mathbf{Y}$  的协方差矩阵,  $\mathbf{C}_{xy} = \mathbf{X}\mathbf{Y}^T$  为互协方差矩阵. 一般假设  $\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x = 1$ ,  $\mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y = 1$ , 上述问题可转化为约束下的最优化问题, 即

$$\begin{aligned} & \max \quad \mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y \\ & \text{s.t.} \quad \begin{cases} \mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x = 1, \\ \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y = 1. \end{cases} \end{aligned} \quad (10)$$

计算时构造拉格朗日等式, 并将问题转化为变换矩阵  $\mathbf{W}_x \in \mathbb{R}^{p \times l}$  和  $\mathbf{W}_y \in \mathbb{R}^{q \times l}$  的特征值方程求解<sup>1)</sup>问题. 细节可参考文献 [32]. 可得,  $\mathbf{Z}_x = \mathbf{W}_x^T \mathbf{X} \in \mathbb{R}^{l \times n}$ ,  $\mathbf{Z}_y = \mathbf{W}_y^T \mathbf{Y} \in \mathbb{R}^{l \times n}$  为变换后的特征, 其中  $l \leq \min\{p, q\}$ . 融合特征为  $\mathbf{Z} = [\mathbf{Z}_x, \mathbf{Z}_y] \in \mathbb{R}^{2l \times n}$ .

### 4.2 核典型相关分析

CCA 是线性运算, 对某些存在非线性关系的特征集合相关效果不好, 而核典型相关分析 (KCCA) 可以通过引入核函数对原始特征进行非线性扩展, 在新的特征空间求解非线性相关关系 [25, 33].

用  $\phi_x$  和  $\phi_y$  分别表示两种特征的非线性映射函数, 对应的映射向量为  $\phi_x(\mathbf{x})$  和  $\phi_y(\mathbf{y})$ . 以  $\mathbf{x}$  为例, 其核变换可以表示为

$$k_x(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi_x(\mathbf{x}_i), \phi_x(\mathbf{x}_j) \rangle, \quad (11)$$

其中,  $\langle \cdot, \cdot \rangle$  表示内积运算. 对应的核矩阵表示为  $\mathbf{K}_x = \mathbf{X}^T \mathbf{X} = [k_x(\mathbf{x}_i, \mathbf{x}_j)]_{i,j=1}^n \in \mathbb{R}^{n \times n}$  和  $\mathbf{K}_y = \mathbf{Y}^T \mathbf{Y} = [k_y(\mathbf{y}_i, \mathbf{y}_j)]_{i,j=1}^n \in \mathbb{R}^{n \times n}$ , KCCA 优化目标是最大化式 (12) 的准则函数 [34]:

$$J(\alpha, \beta) = \frac{\mathbf{w}_x^T \mathbf{K}_x \mathbf{K}_y \mathbf{w}_y}{\sqrt{(\mathbf{w}_x^T \mathbf{K}_x \mathbf{K}_x \mathbf{w}_x + \kappa \mathbf{w}_x^T \mathbf{K}_x \mathbf{w}_x) \cdot (\mathbf{w}_y^T \mathbf{K}_y \mathbf{K}_y \mathbf{w}_y + \kappa \mathbf{w}_y^T \mathbf{K}_y \mathbf{w}_y)}}, \quad (12)$$

其中,  $\mathbf{w}_x \in \mathbb{R}^n$ ,  $\mathbf{w}_y \in \mathbb{R}^n$ ,  $0 < \kappa < 1$  是正则化参数. 在求解时, 通常基于部分 Gram-Schmidt 正交 (PGSO) 算法 [32], 求解出对应的变换矩阵  $\mathbf{W}_x \in \mathbb{R}^{n \times l}$  和  $\mathbf{W}_y \in \mathbb{R}^{n \times l}$ , 得到特征  $\mathbf{Z}_x = \mathbf{K}_x \mathbf{W}_x \in \mathbb{R}^{l \times n}$  和  $\mathbf{Z}_y = \mathbf{K}_y \mathbf{W}_y \in \mathbb{R}^{l \times n}$ .

1) 在特征值求解过程中, 保留非零特征值对应的特征向量.

### 4.3 判别核典型相关分析

判别核典型相关分析 (discriminant kernel canonical correlation analysis, DKCCA) 是在特征融合过程中增加了特征的类别标签信息. DKCCA 首先对特征数据进行非线性映射,  $\Gamma: x \rightarrow K$ ; 然后进行判别相关分析 (DCA) [35], 通过将特征约束转换为对特征散布矩阵的优化问题, 保证类内特征距离尽可能缩小, 类间距离特征尽可能扩大; 最后基于 CCA 完成特征融合.

以两种特征  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$  和  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n] \in \mathbb{R}^{q \times n}$  为例, 阐述算法的计算过程.

首先, 进行非线性映射, 方法同上述 KCCA 的映射部分 [25]. 在具体实现时, 需要基于奇异值分解对核矩阵  $\mathbf{K}_x$  与  $\mathbf{K}_y$  进行近似处理, 提取其主成分, 即  $\mathbf{K}_x = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ , 其中,  $\mathbf{\Sigma}$  是对角阵, 对角元素为奇异值,  $\mathbf{U}$  和  $\mathbf{V}$  都是酉矩阵. 选取大于门限值  $\eta$  的  $d$  个奇异值构造新的奇异值矩阵  $\mathbf{\Sigma}_d$  和相应的  $\mathbf{U}_d$  和  $\mathbf{V}_d$ , 得到

$$\mathbf{K}_x \approx \mathbf{U}_d \mathbf{\Sigma}_d \mathbf{V}_d^T. \quad (13)$$

同理可得  $\mathbf{K}_y$ . 为方便表述, 核变换后的特征仍用符号  $\mathbf{X}$  和  $\mathbf{Y}$  表示.

其次, 考虑特征的标签信息. 假定有  $c$  个同型号的辐射源个体, 即有  $c$  类原始指纹特征数据, 共有  $n$  个样本,  $n = \sum_{i=1}^c n_i$ , 其中  $n_i$  为第  $i$  类的样本数目. 以  $\mathbf{X}$  特征为例, 令  $\mathbf{x}_{ij} \in \mathbf{X}$  表示第  $i$  个辐射源的第  $j$  个  $\mathbf{X}$  特征样本,  $\bar{\mathbf{x}}_i$  为第  $i$  类的类心  $\bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ij}$ ; 整体的均值表示为  $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^{n_i} \mathbf{x}_{ij} = \frac{1}{n} \sum_{i=1}^c n_i \bar{\mathbf{x}}_i$ . 特征  $\mathbf{X}$  的类间散布矩阵定义为

$$\mathbf{S}_{bx} = \sum_{i=1}^c n_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T = \mathbf{\Phi}_{bx} \mathbf{\Phi}_{bx}^T, \quad (14)$$

其中,  $\mathbf{\Phi}_{bx} = [\sqrt{n_1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}), \sqrt{n_2}(\bar{\mathbf{x}}_2 - \bar{\mathbf{x}}), \dots, \sqrt{n_c}(\bar{\mathbf{x}}_c - \bar{\mathbf{x}})]$ ,  $\mathbf{\Phi}_{bx} \in \mathbb{R}^{p \times c}$ ,  $\mathbf{S}_{bx} \in \mathbb{R}^{p \times p}$ .

当  $\mathbf{S}_{bx}$  为对角矩阵时, 数据的可分性最强. 因此, 结合式 (14), 上述问题转变为寻找一个变换矩阵将  $\mathbf{\Phi}_{bx} \mathbf{\Phi}_{bx}^T$  对角化. 考虑到矩阵的维度情况,  $\mathbf{\Phi}_{bx} \mathbf{\Phi}_{bx}^T \in \mathbb{R}^{p \times p}$ , 而  $\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx} \in \mathbb{R}^{c \times c}$ ,  $p \gg c$ , 由于两者的特征向量存在一定的映射关系 [36], 因此对  $\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}$  进行对角化处理更加便捷. 对角化的过程主要有 3 步:

第 1 步, 对  $\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}$  进行特征值分解, 即求解

$$\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx} \mathbf{v}_i = \mu_i \mathbf{v}_i, \quad (15)$$

$$\mathbf{\Phi}_{bx} \mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx} \mathbf{v}_i = \mu_i \mathbf{\Phi}_{bx} \mathbf{v}_i, \quad (16)$$

其中,  $\mathbf{v}_i$  表示  $\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}$  的特征向量,  $\mu_i$  表示对应的特征值, 可知  $\mathbf{\Phi}_{bx} \mathbf{v}_i$  是  $\mathbf{\Phi}_{bx} \mathbf{\Phi}_{bx}^T$  的特征向量, 两者特征值保持一致.

将特征向量  $\mathbf{v}_i$  组成矩阵  $\mathbf{P}$ , 有

$$\mathbf{P}^T (\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}) \mathbf{P} = \hat{\mathbf{\Lambda}}, \quad (17)$$

由于  $\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}$  对称半正定,  $\mathbf{P}$  是正交矩阵,  $\hat{\mathbf{\Lambda}}$  是实非负特征值以降序排列的对角矩阵.

第 2 步, 选择  $\mathbf{P}$  的前  $r$  个特征向量组成矩阵  $\mathbf{Q}^2$ , 从而  $\mathbf{Q}^T (\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}) \mathbf{Q} = \hat{\mathbf{\Lambda}}_{(r \times r)}$ . 令  $\hat{\mathbf{W}}_{bx} = \mathbf{Q} \hat{\mathbf{\Lambda}}^{-1/2}$ , 有  $\hat{\mathbf{W}}_{bx}^T (\mathbf{\Phi}_{bx}^T \mathbf{\Phi}_{bx}) \hat{\mathbf{W}}_{bx} = \mathbf{I}$ , 其中,  $\mathbf{I}$  为单位矩阵. 为方便表述, 不同大小的单位矩阵均用  $\mathbf{I}$  表示. 令  $\mathbf{Q} \rightarrow \mathbf{\Phi}_{bx} \mathbf{Q}$ , 得到

$$(\mathbf{\Phi}_{bx} \mathbf{Q})^T \mathbf{S}_{bx} (\mathbf{\Phi}_{bx} \mathbf{Q}) = \hat{\mathbf{\Lambda}}^2 = \mathbf{\Lambda}, \quad (18)$$

2) 一般为非零特征值对应的前  $r$  个特征向量.



其中,  $\Lambda \in \mathbb{R}^{r \times n}$ .

第 3 步, 构造变换矩阵  $\mathbf{W}_{bx} = \Phi_{bx} \mathbf{Q} \hat{\Lambda}^{-1/2}$ , 使得  $\mathbf{W}_{bx}^T \mathbf{S}_{bx} \mathbf{W}_{bx} = \mathbf{I}$ .  $\mathbf{W}_{bx}$  即为要找的  $\mathbf{S}_{bx}$  的对角化变换矩阵.

利用该变换将特征  $\mathbf{X}$  映射到同一空间, 得到

$$\mathbf{Z}'_x = \mathbf{W}_{bx}^T \mathbf{X}, \quad (19)$$

其中,  $\mathbf{Z}'_x \in \mathbb{R}^{r \times n}$  为  $\mathbf{X}$  变换后的新的特征. 同理, 对特征  $\mathbf{Y}$ , 有  $\mathbf{W}_{by}^T \mathbf{S}_{by} \mathbf{W}_{by} = \mathbf{I}$  与  $\mathbf{Z}'_y = \mathbf{W}_{by}^T \mathbf{Y}$ .

最后, 参照 CCA, 将更新后的特征  $\mathbf{Z}'_x$  与  $\mathbf{Z}'_y$  进行融合, 即对角化特征间协方差矩阵  $\mathbf{S}'_{xy} = \mathbf{Z}'_x \mathbf{Z}'_y{}^T \in \mathbb{R}^{r \times r}$ . 结合奇异值分解有

$$\mathbf{S}'_{xy} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \Rightarrow \mathbf{U}^T \mathbf{S}'_{xy} \mathbf{V} = \mathbf{\Sigma}. \quad (20)$$

令  $\mathbf{W}_{cx} = \mathbf{U} \mathbf{\Sigma}^{-1/2}$ ,  $\mathbf{W}_{cy} = \mathbf{V} \mathbf{\Sigma}^{-1/2}$ , 可以得到  $(\mathbf{U} \mathbf{\Sigma}^{-1/2})^T \mathbf{S}'_{xy} (\mathbf{V} \mathbf{\Sigma}^{-1/2}) = \mathbf{I}$ , 故最终的特征集合为<sup>[35]</sup>

$$\begin{aligned} \mathbf{Z}^*_x &= \mathbf{W}_{cx}^T \mathbf{Z}'_x = \mathbf{W}_{cx}^T \mathbf{W}_{bx}^T \mathbf{X} = \mathbf{W}_x \mathbf{X}, \\ \mathbf{Z}^*_y &= \mathbf{W}_{cy}^T \mathbf{Z}'_y = \mathbf{W}_{cy}^T \mathbf{W}_{by}^T \mathbf{Y} = \mathbf{W}_y \mathbf{Y}, \end{aligned} \quad (21)$$

其中,  $\mathbf{W}_x = \mathbf{W}_{cx}^T \mathbf{W}_{bx}^T$  和  $\mathbf{W}_y = \mathbf{W}_{cy}^T \mathbf{W}_{by}^T$  是最终的变换矩阵,  $\mathbf{Z}^*_x, \mathbf{Z}^*_y \in \mathbb{R}^{l \times n}$ . 对应地,  $\mathbf{Z}^*_x$  的类间散布矩阵可以表示为

$$\mathbf{S}^*_{bx} = \mathbf{W}_{cx}^T \mathbf{W}_{bx}^T \mathbf{S}_{bx} \mathbf{W}_{bx} \mathbf{W}_{cx}. \quad (22)$$

由上文可知,  $\mathbf{W}_{bx}^T \mathbf{S}_{bx} \mathbf{W}_{bx} = \mathbf{I}$  和  $\mathbf{U}$  都是正交矩阵, 因此有

$$\mathbf{S}^*_{bx} = (\mathbf{U} \mathbf{\Sigma}^{-\frac{1}{2}})^T (\mathbf{U} \mathbf{\Sigma}^{-\frac{1}{2}}) = \mathbf{\Sigma}^{-1}, \quad (23)$$

即  $\mathbf{S}^*_{bx} = \mathbf{\Sigma}^{-1}$  是对角矩阵.

#### 4.4 多域特征融合策略

多域特征融合 (MDKCCA) 的基本思路是基于 DKCCA 分别进行不同特征组的融合. 计算顺序上首先融合同一特征域的多个特征, 然后进行不同域特征融合. 对每次 DKCCA 融合而言, 需要先通过训练求解出最终的变换矩阵  $\mathbf{W}$ , 而后续识别过程中可以依据训练得到变换矩阵直接运算. 最后, 将式 (21) 中两种 DKCCA 特征  $\mathbf{Z}^*_x$  和  $\mathbf{Z}^*_y$  以直接级联的方式实现融合.

以 3 个域 6 种特征为例, 融合的流程如图 5 所示. 其中特征域为  $\mathbf{A}, \mathbf{B}, \mathbf{C}$ , 每个特征域内有两个特征分别为  $A_1, A_2, B_1, B_2, C_1, C_2$ . 融合特征结果用对应的域标签表示  $(A, B, AB, \dots)$ , 3 种域的最终融合结果用特征  $ABC$  表示. 特征域  $\mathbf{A}$  内融合特征  $A$  可以表示为  $A = \mathbf{W}_{A1} \mathbf{K}_{A1} \oplus \mathbf{W}_{A2} \mathbf{K}_{A2}$ , 其中,  $\oplus$  表示级联运算,  $\mathbf{W}_m$  和  $\mathbf{K}_m$  ( $m = A_1, A_2$ ) 分别表示对应 DKCCA 的变换矩阵和核矩阵. 同理, 融合特征  $AB$  可以表示为

$$AB = \mathbf{W}_A \mathbf{K}_A \oplus \mathbf{W}_B \mathbf{K}_B = (\mathbf{W}_A (\mathbf{W}_{A1} \mathbf{K}_{A1} \oplus \mathbf{W}_{A2} \mathbf{K}_{A2})) \oplus \mathbf{W}_B (\mathbf{W}_{B1} \mathbf{K}_{B1} \oplus \mathbf{W}_{B2} \mathbf{K}_{B2}). \quad (24)$$

3 种特征域的最终融合特征  $ABC$  可以表示为

$$ABC = \mathbf{W}_{AB} \mathbf{K}_{AB} \oplus \mathbf{W}_C \mathbf{K}_C. \quad (25)$$

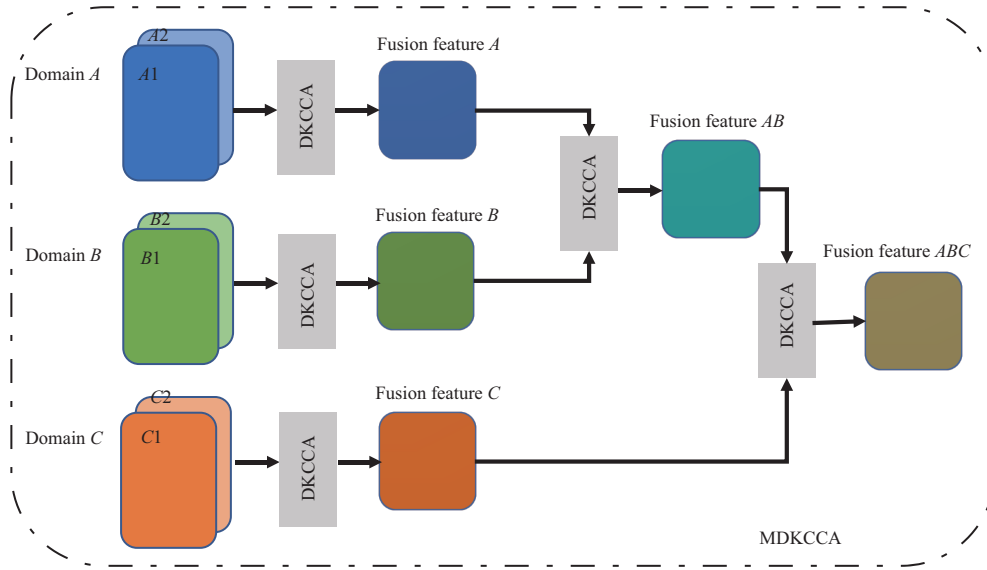


图 5 (网络版彩图) MDKCCA 特征融合策略示意图 (以 3 种域特征为例)

Figure 5 (Color online) Schematic diagram of the MDKCCA feature fusion strategy (taking features of three domains as example)

更一般地, 第  $i$  次 DKCCA 融合的结果  $M_i$  可以表示为

$$M_i = W_{i-1}^1 K_{i-1}^1 \oplus W_{i-1}^2 K_{i-1}^2, \quad (26)$$

其中,  $K_{i-1}^1, K_{i-1}^2$  分别表示用于第  $i$  次融合的两个特征  $M_{i-1}^1$  和  $M_{i-1}^2$  的核矩阵,  $W_{i-1}^1, W_{i-1}^2$  为对应的变换矩阵.  $M_0, W_0, K_0$  为原始未融合的特征及其变换矩阵和核矩阵.

## 5 实验结果与分析

### 5.1 数据集与参数设置

为测试本文算法对不同体制信号的不同辐射源设备的识别性能, 基于 4 种类型的实际采集数据集进行实验, 包括基于 802.11 系列协议的无线网卡信号 IEEE 802.11b 和 IEEE 802.11g, OFDM-16QAM 体制的电台通信信号 (radio) 以及二次雷达应答信号 (secondary surveillance radar, SSR), 共计 35 个辐射源个体 11977 个数据样本. 上述数据集分别用 11b, 11g, radio 和 SSR 表示, 详细情况如表 1 所示.

本文重点研究特征融合方法, 因此在识别环节选用结构简单的 KNN (k-nearest neighbor) 分类器, 将识别准确率作为本文方法的评价指标. 其中, KNN 分类器的最近邻个数为 5. 若非特别说明, 其他参数设置如下: 随机选择数据的 50% 作为训练集, 其余 50% 作为测试集; 识别正确率数据均为 200 次 Monte Carlo 的平均结果; 选择高斯核函数作为非线性映射的核方法.

### 5.2 结果分析

#### 5.2.1 原始特征性能

所选的 4 个特征域 8 种原始特征的维度情况如表 2 所示. 为便于对比, 此处对特征的维度进行了

表 1 数据集细节描述  
Table 1 Description of datasets

	Dataset	Signal	Emitter model	Device number	Sample number
1	11b	IEEE 802.11b	Tenda U3	5	2500
2	11g	IEEE 802.11g	TPLink WN322G+	5	2500
3	Radio	OFDM-16QAM	Anykey AKDS700	3	710
4	SSR	Secondary radar signal	-	20	6267
Total	-	-	-	35	11977

表 2 原始特征维度  
Table 2 Dimension of original feature

Feature	Domain T		Domain F		Domain S		Domain B		Total
	IA	Env	IF	IP	Slice1	Slice2	DIB	SIB	
Dimension	100	100	120	120	121	121	256	128	1026

专门设置: 不同数据集上同一种特征的维数保持一致; 由于不同域特征超参数设置差异, 如滑窗宽度、傅里叶变换点数等, 导致不同域特征的维度不同; 同一特征域内不同特征的维度保持一致. 由于双谱的积分路径存在差异, 因此在双谱域中两种特征维度不相同.

8 种原始特征在 4 种数据集的识别性能如图 6 所示, 图中不同的条带分别代表不同的特征的识别结果, 从左向右分别按顺序对应图中图例 (下文相同), 每个数据集的最后一根条带为 8 种特征的平均识别率. 从图 6 中可以得到如下结论:

第一, 对单个数据集来说, 不同特征的性能优劣不一, 甚至存在较大差异. 如, 对数据集 SSR 而言, 模糊域特征均能达到 90% 以上, 而双谱域特征效果较差, 只能达到 30% 左右.

第二, 同一种特征方法对不同类型信号的适应能力不同. 例如特征 Slice2 在数据集 radio 上识别率大于 95%, 可以作为最优特征, 而在数据集 11g 上仅有 70%.

实验结果证明了不同域指纹特征性能存在差异, 印证了上文有关不同域特征间存在互补与新息的分析.

### 5.2.2 特征融合性能

本文算法融合后特征维度情况如表 3 所示<sup>3)</sup>, 单一域内两种特征融合后的识别效果如图 7(a) 所示, 多域融合结果如图 7(b) 所示.

(1) **单域融合.** 对比图 6 和 7(a) 可知, 单域内特征融合后性能优劣排序基本没有明显变化. 例如, 对数据集 SSR 而言, 相比单一特征, 融合后的双谱域特征的识别性能得到了提升, 由 30% 左右增长到接近 60%, 但是仍然远远低于最优特征 (模糊域特征), 说明了单域内特征融合带来的新息有限.

(2) **多域融合.** 图 7(b) 展示了多域特征融合的识别结果, 其中每个数据集的 3 个值分别对应时域与时频域融合结果 (T+F)、双谱域与模糊域融合结果 (S+B), 以及 4 种域融合结果 (TF+SB). 多域融合后识别率在 90% 左右, 其中 4 种域融合的结果 (图 7(b) 标注值) 高于两种域融合结果, 能够达到 95% 以上, 在数据集 radio 和 SSR 更是超过 99%. 与图 7(a) 对比, 不同域特征融合提升效果明

3) 考虑到不同数据集的目标个数、数据规模, 以及特征的分布情况并不相同, 各数据集在经过融合后的特征维度未保持一致. 其中数据集 SSR 中包含 20 个目标, 特征维度数最大, 为 38. 下文对特征维度与识别效果之间的关系进行了专门探究.

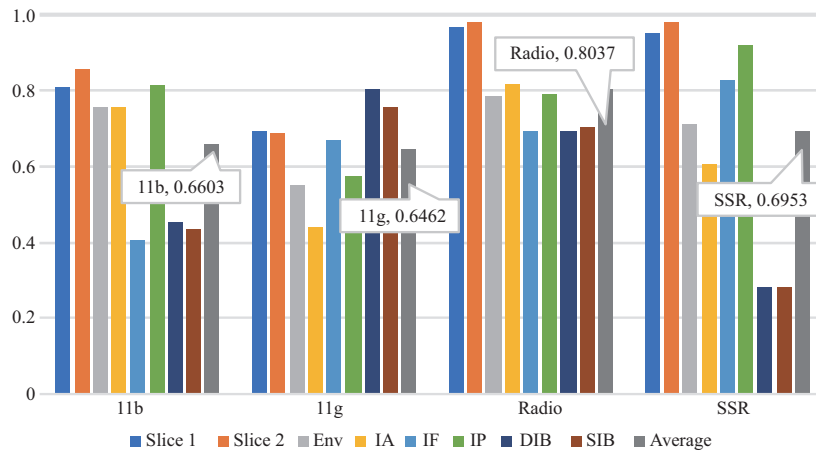


图 6 原始特征的识别结果

Figure 6 Identification results of original features

表 3 MDKCCA 特征维度

Table 3 Feature dimension of MDKCCA

Feature	Dataset			
	11b	11g	Radio	SSR
Dimension	8	8	4	38

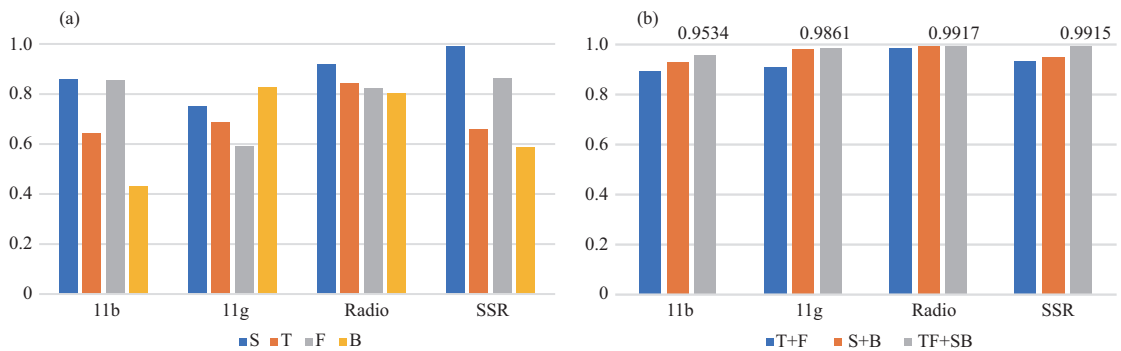


图 7 特征融合结果

Figure 7 Identification results of feature fusion. (a) Single-domain fusion; (b) multi-domain fusion

显高于单域内的特征融合. 而且识别性能随着融合特征域数目的增加而进一步提升, 且均优于最优单一特征. 这是因为不同特征域之间存在大量新息, 4 种域特征融合为辐射源指纹的观测提供了更多的角度.

图 8 给出了最终 4 种域特征经本文算法融合后在 4 种数据集上的混淆矩阵. 绝大多数的辐射源个体的正确识别率 (对角线元素) 均能达到 90% 以上.

(3) 特征维度影响分析. 除识别性能提升外, 对比表 2 中的原始特征维度, 本文算法的特征维度大幅下降, 从而大大降低了对分类器的要求. 数据集 SSR 的最终特征为 38 维, 在目标数为 20 的情况下识别率能达到 99.15%; 数据集 11b 和 11g 的特征均为八维, 性能提升较为明显, 相比于最优原始特征

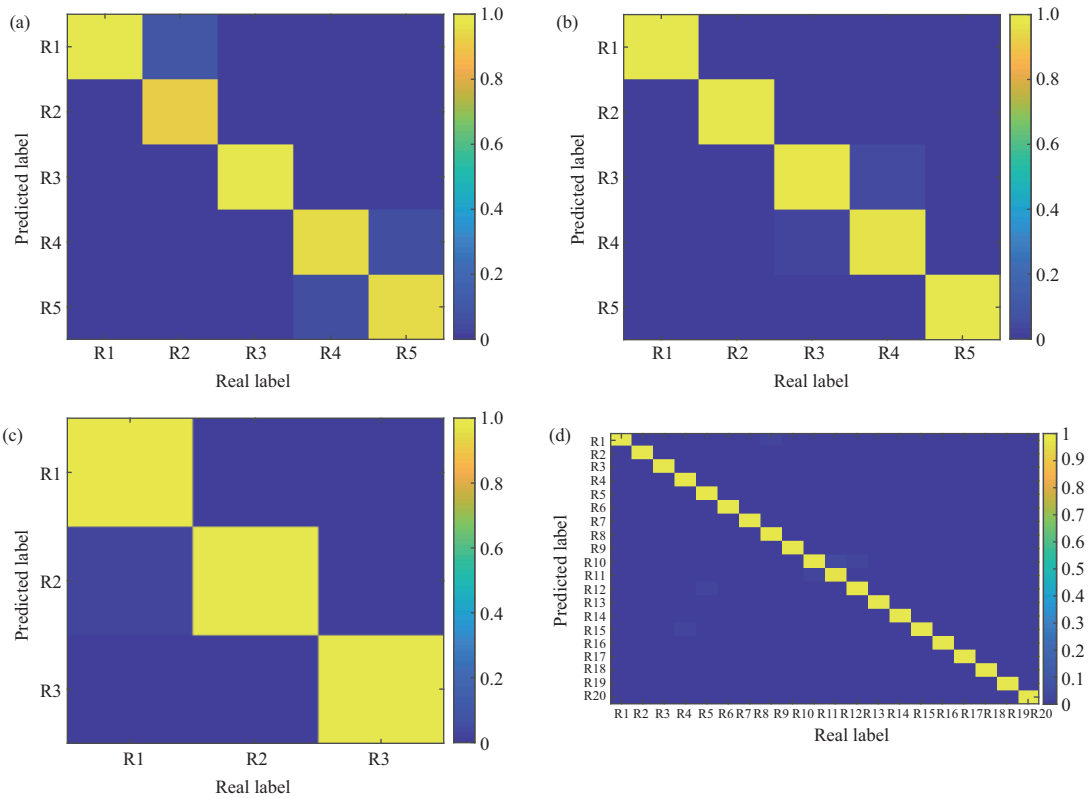


图 8 MDKCCA 特征识别混淆矩阵

Figure 8 Recognition confusion matrix of MDKCCA feature. (a) 11b; (b) 11g; (c) radio; (d) SSR

提高幅度在 10% 以上, 最终识别结果均能达到 95% 以上; 数据集 radio 融合特征维度最低, 仅为四维, 仍能够达到 99% 的正确识别率.

### 5.2.3 参数影响

本小节探究本文算法中超参数设置对识别效果的影响, 包括核变换过程中正则化参数  $\kappa$ 、奇异值分解中的门限值  $\eta$ , 以及特征维度  $l$ .  $\kappa$  和  $\eta$  的分析主要在数据集 11b 上进行, 特征维度  $l$  的分析在数据集 SSR 上完成. 这是因为数据集 11b 中单个目标的数据量较多且相对均衡, 而数据集 SSR 中包含的目标个数最多, 特征维度最高, 因此在该数据集 SSR 上探究特征维数对识别效果的影响. 本文算法的识别性能随不同超参数变化的曲线如图 8 所示.

(1) 正则化参数影响. 从图 9(a) 可以看出,  $\kappa$  对识别率的影响有限, 识别率保持在 95% 左右轻微波动. 当  $\kappa = 0.7$  时, 识别效果最好.

(2) 门限值影响. 利用奇异值分解对核矩阵进行近似处理时需要设置合适的门限值, 即 4.3 小节中的门限值  $\eta$ . 从图 9(b) 可以看出, 门限值大小会影响识别性能, 在  $\eta = 0.1$  时, 识别率最高; 当  $\eta > 0.15$  时, 识别效果出现下降.

(3) 特征维度影响. 由图 9(c) 可知, 识别性能随着特征维度的增加而提升, 维度在 10 以下时 (5~10), 能够达到 90% 以上的识别率. 表明本文算法能在大幅降低特征维度的同时保证识别性能. 当维度超过 14 时, 识别率提升至 97% 左右. 随着维度进一步增加, 识别性能变化不大. 在实际应用中可以根据需求, 确定合适的融合特征维度.

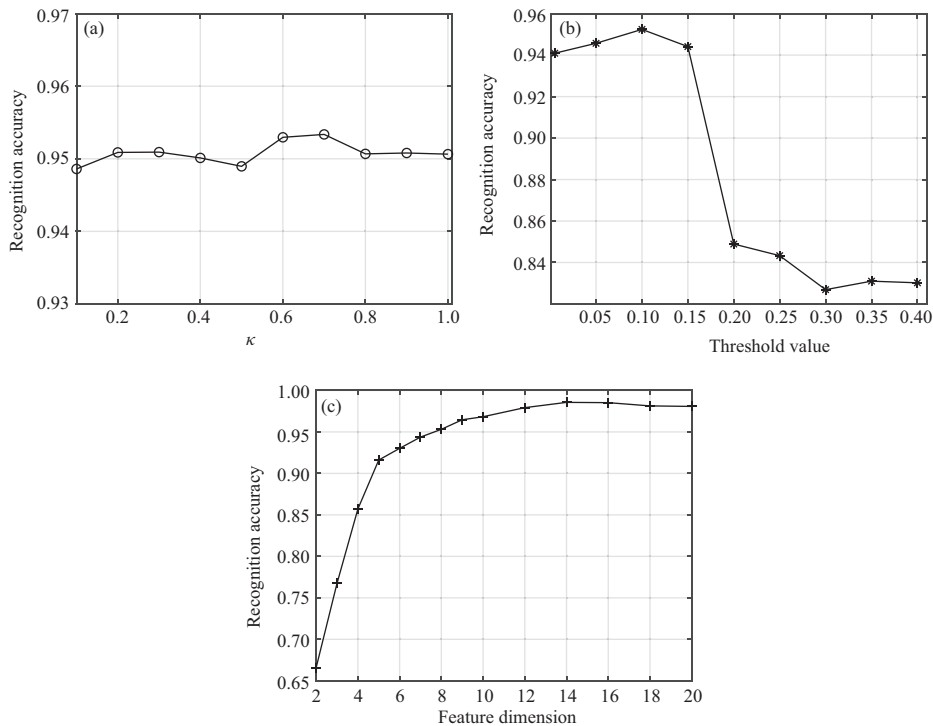


图 9 参数对识别效果的影响

**Figure 9** Influence of parameters on the recognition effect. (a) Regularization parameter; (b) threshold value; (c) feature dimension

### 5.3 性能对比实验

为进一步验证本文算法的性能, 选取 4 类方法进行了比较: 一类是简单特征综合方法, 包括最优单一特征 (optimal feature, OF)、原始特征平均 (average)、直接串联融合方法 (series-connected feature, SF), 以及 PCA 方法; 二类是单域特征融合算法, 包括 DCA 算法<sup>[35]</sup>、CCA 算法<sup>[32]</sup>; 三类是基于深度学习的特征综合算法 (DL)<sup>[23,37,38]</sup>; 四类是多域特征融合算法. 性能结果如表 4 所示.

其中, “最优特征”指 8 种原始特征中识别效果最好的单个特征, 不同数据集的最优特征是不同的, 例如数据集 11b 的最优特征为 Slice2, 数据集 11g 的为 DIB; “平均”表示 8 种原始特征识别率的平均值; 直接串联方法表示直接串联全部 8 种原始特征, 维度为 1026; PCA 方法是直接将串联特征进行 PCA 降维, 降维后的特征维度与本文算法保持一致, 同表 3 所示. “DCA”, “CCA”表示基于 DCA 或 CCA 算法对同一域内两个特征 (单域内) 融合的最佳识别结果. 除基于深度学习算法外, 其余算法的识别结果均基于相同的 5-KNN 分类器和 200 次 Monte Carlo 实验. DL 方法基于卷积神经网络结构, 网络结构参照文献<sup>[37,38]</sup>, 选择用直接串联特征的数据作为网络输入<sup>[23]</sup>, 数据集的 80% 作为训练集, 10% 作为验证集, 剩余 10% 作为测试集, 表 4 的数据为测试集识别率.

(1) 对比简单特征综合方法. 从表 4 可以看出, 串联特征维度最高且效果最差, 在此基础上进行 PCA 分析后, 仅有部分提升, 仍然低于最优特征的性能和特征平均识别率. 可见, 简单特征综合并不能带来性能提升, 因此在 SEI 实际应用中通常需要结合专家经验人工从备选特征中寻找并确定最优特征. 然而对不同类型信号的最优特征并不同, 使得系统在面临新信号和新场景时计算量大. 本文算法性能上明显高于简单特征综合算法以及人工寻优确定的最优特征, 同时融合过程无需利用信号类型

表 4 对比实验结果  
Table 4 Comparative experimental results

Algorithm	Test dataset			
	11b	11g	Radio	SSR
PCA	0.4663	0.6983	0.8120	0.2143
DCA	0.7898(S)	0.8212(B)	0.8554(S)	0.9287(B)
CCA	0.7237(F)	0.7190(B)	0.8140(B)	0.9811(S)
SF	0.4029	0.6697	0.7251	0.2589
DL	0.5100	0.6800	0.6500	0.7750
Average	0.6741	0.6462	0.8037	0.6953
OF	0.8558	0.8042	0.9815	0.9824
Proposed method	<b>0.9534</b>	<b>0.9861</b>	<b>0.9917</b>	<b>0.9915</b>

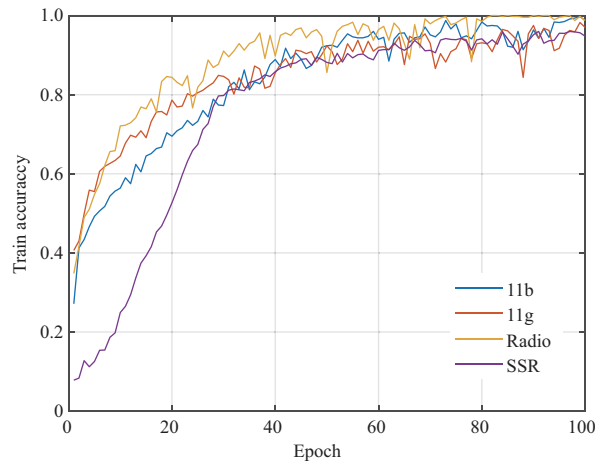


图 10 DL 融合算法训练集识别率

Figure 10 Train accuracy of the DL-based feature fusion method

的先验信息, 减少了相关寻优环节, 降低了 SEI 应用的技术门槛.

(2) 对比神经网络融合方法. 训练数据的识别效果如图 10 所示. 从图中可知, 基于 DL 的融合算法在训练集上识别率大于 80%, 然而在测试集上识别率下降了 20% 左右, 这是由于标注数据有限 (表 1), 平均每个目标的样本总数最多只有 500 左右, 而训练样本均不超过 400. 无法满足数据驱动的 DL 算法的训练需求. 本文算法在训练数据较少的情况下, 鲁棒性优于 DL 算法, 更符合 SEI 实际应用场景.

(3) 对比单域融合方法. 表格标注了最优单域内融合特征对应的域代号, 例如数据集 11b 上 CCA 方法的结果为时频域 (F) 特征融合结果. 由表可知, DCA 的性能总体上高于 CCA, 而在目标数目较多的数据集 SSR 上略低于 CCA. 此类融合算法性能明显高于简单综合方法, 但却不及最优单一特征, 融合效果有待提升. 而且由于单域融合的限制, 在全部数据集上的识别结果均低于本文算法.

(4) 不同多域融合方法分析. 为进一步验证本文采用的非线性扩展和标注信息的有效性, 将算法中最后一个阶段的特征融合过程更换为 DCA 或 CCA 进行比较, 其他环节保持不变. 由于采用了多域融合策略, 因而将基于 DCA 或 CCA 的对比方法分别用 MDCA 和 MCCA 表示, 结果如表 5 所示. 综合表 4 和 5 可以看出, 第一, 多域融合类方法优于单域融合方法, 在数据集 11b, 11g 和 Radio 上的识

表 5 多域策略下的不同方法识别效果

Table 5 Identification results of different methods under multi-domain strategy

Dataset	Fusion method of TF+SB		
	MDCA	MCCA	Proposed method
11b	0.9275	0.9228	0.9534
11g	0.9269	0.9423	0.9861
Radio	0.9099	0.9248	0.9917
SSR	0.7996	0.9740	0.9915

别率提升平均在 10% 左右, 识别率都超过 90%, 且优于最优单域特征; 第二, MDCA, MCCA 提升效果并不稳定, 例如对数据集 SSR 而言, MDCA 的 4 种特征域融合结果 (0.7996) 低于 DCA 最优单一特征域 (表 4 (0.9287)). 第三, 本文算法由于利用了数据标注, 融合后的特征对指纹的表征更加精确, 类间差异性更大, 与 MDCA, MCCA 方法相比, 识别率最高.

## 6 总结

本文针对 SEI 的特征综合利用问题进行了研究, 提出一种基于 MDKCCA 的指纹特征融合方法. 本文算法能够结合标签信息与特征之间的非线性关系, 通过核变换将特征映射到高维空间实现多域特征的有效融合, 大幅提高了特征的利用率, 同时实现特征降维, 提高分类识别的运算效率. 在多种实测数据集上进行了 4 个指纹特征域的 8 种常见特征的融合实验. 结果表明不同特征域的特征之间存在互补性, 识别效果非最佳的特征也可辅助 SEI 算法性能提升, 而且参与融合的特征域种类越多, 特征方法之间新息越明显, 融合效果越好. 本文算法识别率在全部数据集上均达到 95% 以上, 不仅优于人为挑选的单一最优特征, 以及简单特征综合方法、基于神经网络的特征综合方法以及基于 DCA 或 CCA 的特征融合方法. 此外, 在算法实现上, 本文算法不需要人工预处理特征和人工挑选最优特征, 也无需人工设定融合权重, 直接对多域特征进行综合利用, 大幅减少对专家知识的依赖, 具有较高的应用价值.

## 参考文献

- 1 Xie F Y, Wen H, Wu J S, et al. Data augmentation for radio frequency fingerprinting via pseudo-random integration. *IEEE Trans Emerg Top Comput Intell*, 2020, 4: 276–286
- 2 Liu M W, Doherty J F. Nonlinearity estimation for specific emitter identification in multipath channels. *IEEE Trans Inform Forensic Secur*, 2011, 6: 1076–1085
- 3 Ru X H, Gao C, Liu Z, et al. Emitter identification based on the structure of unintentional modulation. In: *Proceedings of the 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2016. 998–1002
- 4 Gok G, Alp Y K, Arikan O. A new method for specific emitter identification with results on real radar measurements. *IEEE Trans Inform Forensic Secur*, 2020, 15: 3335–3346
- 5 Gong J L, Xu X D, Lei Y K. Unsupervised specific emitter identification method using radio-frequency fingerprint embedded InfoGAN. *IEEE Trans Inform Forensic Secur*, 2020, 15: 2898–2913
- 6 Davaslioglu K, Soltani S, Erpek T, et al. DeepWiFi: cognitive WiFi with deep learning. *IEEE Trans Mobile Comput*, 2021, 20: 429–444
- 7 Roy D, Mukherjee T, Chatterjee M, et al. RFAL: adversarial learning for RF transmitter identification and classification. *IEEE Trans Cogn Commun Netw*, 2020, 6: 783–801



- 8 Kenneth I T, Duley P D, Hyatt M H. Specific emitter identification and verification. *Technol Rev J*, 2003, 113: 133
- 9 Fei Z, Huang Y L, Jian C. Specific emitter identification based on linear polynomial fitting of the energy envelope. In: *Proceedings of International Conference on Electronics Information & Emergency Communication*, 2016
- 10 Wang H W, Zhao G Q, Wang Y J. Specific emitter identification based on higher order moment of the envelope's front edge. *Modem Radar*, 2010, 10: 47-50 [王宏伟, 赵国庆, 王玉军. 基于脉冲包络前沿高阶矩特征的辐射源个体识别. *现代雷达*, 2010, 10: 47-50]
- 11 Wu L W, Zhao Y Q, Feng M F, et al. Specific emitter identification using IMF-DNA with a joint feature selection algorithm. *Electronics*, 2019, 8: 934
- 12 Ru X H, Liu Z, Huang Z T, et al. Evaluation of unintentional modulation for pulse compression signals based on spectrum asymmetry. *IET Radar Sonar Nav*, 2017, 11: 656-663
- 13 Seddighi Z, Ahmadzadeh M R, Taban M R. Radar signals classification using energy-time-frequency distribution features. *IET Radar Sonar Nav*, 2020, 14: 707-715
- 14 Digne F, Baussard A, Khenchaf A, et al. Classification of radar pulses in a naval warfare context using Bézier curve modeling of the instantaneous frequency law. *IEEE Trans Aerosp Electron Syst*, 2017, 53: 1469-1480
- 15 Zhou Y P, Wang X, Chen Y, et al. Specific emitter identification via bispectrum-radon transform and hybrid deep model. *Math Problem Eng*, 2020, 2020: 1-17
- 16 Lin Y, Jia J C, Wang S, et al. Wireless device identification based on radio frequency fingerprint features. In: *Proceedings of IEEE International Conference on Communications (ICC)*, 2020
- 17 Wang H H, Zhang T. Extraction algorithm of communication signal characteristics based on improved bispectra and time-domain analysis. *J Signal Process*, 2017, 33: 864-871 [王欢欢, 张涛. 结合时域分析和改进双谱的通信信号特征提取算法. *信号处理*, 2017, 33: 864-871]
- 18 Sankhe K, Belgiovine M, Zhou F, et al. No radio left behind: radio fingerprinting through deep learning of physical-layer hardware impairments. *IEEE Trans Cogn Commun Netw*, 2020, 6: 165-178
- 19 Jian T, Rendon B C, Ojuba E, et al. Deep learning for RF fingerprinting: a massive experimental study. *IEEE Int Thing M*, 2020, 3: 50-57
- 20 Sun L T, Huang Z T, Wang X, et al. Overview of radio frequency fingerprint extraction in specific emitter identification. *J Radar*, 2020, 9: 1014-1031 [孙丽婷, 黄知涛, 王翔, 等. 辐射源指纹特征提取方法述评. *雷达学报*, 2020, 9: 1014-1031]
- 21 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 Trans Inform Forensic Secur*, 2019, 14: 581-591
- 22 Bitar N, Muhammad S, Refai H H. Wireless technology identification using deep convolutional neural networks. In: *Proceedings of the 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017
- 23 Liu Z M. Multi-feature fusion for specific emitter identification via deep ensemble learning. *Digit Signal Process*, 2021, 110: 102939
- 24 Du C. Research on manifold learning and its applications. Dissertation for Ph.D. Degree. Changsha: National University of Defense Technology, 2014 [杜春. 流形学习及其应用算法研究. 博士学位论文. 长沙: 国防科学技术大学, 2014]
- 25 Shi Y, Ji H B. Kernel canonical correlation analysis for specific radar emitter identification. *Electron Lett*, 2014, 50: 1318-1320
- 26 Ye H, Liu Z, Jiang W. Comparison of unintentional frequency and phase modulation features for specific emitter identification. *Electron Lett*, 2012, 48: 875
- 27 Ru X H, Liu Z, Jiang W L, et al. Recognition performance analysis of instantaneous phase and its transformed features for radar emitter identification. *IET Radar Sonar Nav*, 2016, 10: 945-952
- 28 Li L, Ji H B. Specific emitter identification based on ambiguity function. *J Electron Inform Technol*, 2009, 31: 2546-2551 [李林, 姬红兵. 基于模糊函数的雷达辐射源个体识别. *电子与信息学报*, 2009, 31: 2546-2551]
- 29 Zhu M Z, Feng Z P, Zhou X D, et al. Specific emitter identification based on synchrosqueezing transform for civil radar. *Electronics*, 2020, 9: 658
- 30 Wang L, Ji H B, Shi Y. Moving radar emitter recognition based on representative-cut feature of ambiguity function. *Syst Eng Electron*, 2010, 8: 1630-1634 [王磊, 姬红兵, 史亚. 基于模糊函数代表性切片的运动雷达辐射源识别. 系

- 统工程与电子技术, 2010, 8: 1630–1634]
- 31 Guan J P. Study of individual transmitter identification based on bispectrum. Dissertation for Master Degree. Xi'an: Xidian University, 2014 [官建平. 基于双谱的辐射源个体识别技术. 硕士学位论文. 西安: 西安电子科技大学, 2014]
  - 32 Hardoon D R, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: an overview with application to learning methods. *Neural Comput*, 2004, 16: 2639–2664
  - 33 Lisanti G, Karaman S, Masi I. Multichannel-kernel canonical correlation analysis for cross-view person reidentification. *ACM Trans Multimed Comput Commun Appl*, 2017, 13: 1–19
  - 34 Bu Y, Lu L L. Speaker recognition based on multi-features fused by deep learning and kernel canonical correlation analysis. *Comput Digit Eng*, 2019, 47: 2185–2189 [卜禹, 陆璐璐. 基于深度学习及核典型相关分析的多特征融合说话人识别. *计算机与数字工程* 2019, 47: 2185–2189]
  - 35 Haghghat M, Abdel-Mottaleb M, Alhalabi W. Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition. *IEEE Trans Inform Forensic Secur*, 2016, 11: 1984–1996
  - 36 Turk M, Pentland A. Eigenfaces for recognition. *J Cogn Neurosci*, 1991, 3: 71–86
  - 37 Merchant K, Revay S, Stantchev G, et al. Deep learning for RF device fingerprinting in cognitive communication networks. *IEEE J Sel Top Signal Process*, 2018, 12: 160–167
  - 38 Yu J B, Hu A Q, Li G Y, et al. A multi-sampling convolutional neural network-based rf fingerprinting approach for low-power devices. In: *Proceedings of Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2019. 1–6

## Radio frequency fingerprinting feature fusion based on multi-domain discriminant kernel canonical correlation analysis

Liting SUN, Xiang WANG\* & Zhitao HUANG

*Department of Electronic Science, National University of Defense Technology, Changsha 410073, China*

\* Corresponding author. E-mail: christopherwx@163.com

**Abstract** Specific emitter identification (SEI) refers to the precise identification of a specific transmitter through the extraction of hardware fingerprint information from the received signal. The majority of related research focuses on the definition and extraction of new fingerprint features, with less emphasis on the comprehensive use of existing features. Because the characteristics of different analysis domains are complementary to the description of the radio frequency fingerprint, this paper proposes a multi-domain feature fusion strategy for SEI based on multi-domain discriminant kernel canonical correlation analysis (MDKCCA), which fully exploits the complementarity between the features of different domains and combines feature tag information. MDKCCA affords multi-domain feature dimensionality reduction and fusion in a high-dimensional space. The algorithm's performance is validated on four different types of real-world data sets using eight common fingerprint features in four feature analysis domains. The results show that this method eliminates the need for manual feature optimization and can significantly reduce the dimensionality of fusion features. The recognition rate on various targets exceeds 95%, which is higher than the best single feature. It is also superior to the simple feature synthesis method based on direct cascade or PCA dimensionality reduction transformation, the neural network-based feature synthesis method, and the feature fusion method based on discriminant canonical correlation methods.

**Keywords** specific emitter identification (SEI), feature fusion, multi-domain radio frequency fingerprint feature, canonical correlation analysis (CCA), feature extraction