

对海雷达目标识别中全极化 HRRP 的特征提取与选择

范学满* 胡生亮 贺静波
(海军工程大学电子工程学院 武汉 430033)

摘 要: 充分、有效地利用目标全极化 HRRP 的特征信息是提高对海雷达目标识别率的研究热点之一。该文利用 CST 软件仿真建立了 7 类海上目标在不同方位角下的全极化 HRRP 数据库。在此基础上, 提取了 4 类共 39 个特征。提出一种基于归一化互信息(NMI)并利用模拟退火(SA)算法进行优化的全局最优特征选择算法, 并命名为 NMI-SA。基于 HRRP 数据集以及 9 个 UCI 数据集, 利用 k -近邻分类器将该算法与另外 3 种常用的特征选择算法进行对比, 结果表明新算法选择的特征具有良好的可分性和较低的冗余度, 最终用于分类时的正确率总体优于其余 3 种算法。最后, 用该算法对全极化 HRRP 的 39 个特征进行重点分析, 选择出 25 个辨别力强、冗余度低的特征。
关键词: 全极化 HRRP; 特征提取; 特征选择; 互信息; 模拟退火

中图分类号: TN959.72

文献标识码: A

文章编号: 1009-5896(2016)12-3261-08

DOI: 10.11999/JEIT160722

Feature Extraction and Selection of Full Polarization HRRP in Target Recognition Process of Maritime Surveillance Radar

FAN Xueman HU Shengliang HE Jingbo

(Institute of Electronics Engineering, Naval University of Engineering, Wuhan 430033, China)

Abstract: Making full and effective use of target polarization information from High Resolution Range Profile (HRRP) is a hot issue for improving the recognition performance of maritime surveillance radar. A HRRP database with seven maritime targets classes from various aspect angles is established, on which thirty-nine features from four categories are defined. A novel feature selection method based on the Normalized Mutual Information (NMI) and Simulated Annealing (SA) algorithm is presented, named as NMI-SA. The effectiveness of the NMI-SA is proved by comparison with three other methods using HRRP dataset and eight from UCI machine learning repository. Finally, the NMI-SA is applied to the HRRP dataset to find twenty-five high discriminant and low redundancy features.

Key words: Fully polarized HRRP; Feature extraction; Feature selection; Mutual information; Simulated annealing

1 引言

研究如何从众多舰艇以及角反射体组成的多目标群中识别出目标舰艇, 已成为对海雷达的一项重要课题。全极化 HRRP 能够提供比单极化 HRRP 更为丰富的目标结构信息, 随着全极化 HRRP 技术在对海雷达上的逐步应用, 基于全极化 HRRP 的特征提取与目标识别技术引起广泛关注^[1]。文献[2,3]将 4 种极化方式下的 HRRP 看作 4 个独立通道分别进行特征提取, 均没有利用极化特征^[4]。文献[5-8]利用极化分解理论提取 $H/A/\alpha/P$ 、极化不变量、

相似性参数等极化特征进行分类识别, 取得了较好的分类效果。本文综合上述单通道 HRRP 特征和极化特征两类特征, 从物理结构、散射随机性、散射矩阵结构相似性、Mueller 矩阵相似性 4 个方面提取了共计 39 个特征, 旨在充分利用全极化 HRRP 各类信息, 更加全面地刻画目标特性。

理论上讲, 用于分类的特征数越多, 各类目标的可分性越强。然而实际并非总是如此, 因为并不是所有特征对分类都起积极作用, 同时维数过高可能引起维数灾难^[9]。本文提取的 39 个特征的辨别力不尽相同, 因此需要进行特征选择, 剔除与分类任务不相关的特征, 从而提高泛化能力^[10,11]。由于互信息(Mutual Information, MI)能够有效反映变量之间的非线性关系且具有十分成熟的理论基础, 因此成为特征选择算法中用于评估特征重要性的最常用准则之一。文献[12]中总结了近 20 年来基于 MI 的

收稿日期: 2016-07-07; 改回日期: 2016-11-01; 网络出版: 2016-12-02

*通信作者: 范学满 oucfanxm@163.com

基金项目: 国家自然科学基金(61401493), 国家部委基金(9140A01010415JB11002)

Foundation Items: The National Natural Science Foundation of China (61401493), The National Ministries Foundation of China (9140A01010415JB11002)

特征选择算法的研究进展,列举了 17 种基于 MI 的特征选择算法,这些算法都是基于贪婪策略递增地进行特征选择,即每步只选择一个使分类相关性最大、特征子集冗余度最小的特征,已经选择的特征不能被剔除,直到所选特征数目满足要求为止,显然这些基于 MI 的特征选择算法都是次优的。

针对上述问题本文利用归一化互信息(Normalized Mutual Information, NMI)代替互信息进行特征间冗余度以及特征(集)与分类任务间相关性的度量,并利用模拟退火算法(Simulated Annealing, SA)^[13]进行优化搜索,从而较为高效地选择出全局最优的特征子集。

2 全极化 HRRP 特征提取

目前,对海雷达的分类识别重点关注舰艇和角反射体目标,本文在角反射体方面选择由 20 个三面角反射体组成的异型角反射体(三面角反射体的垂直边长为 1.52 m),取 6 个异型角反射体两两间隔 30 m 构成角反射体阵列,如图 1 所示。在舰艇方面选择主要结构参数如表 1 所示的 6 艘舰艇。

2.1 全极化 HRRP 数据库建立

由于上述舰艇目标多为非合作目标,因此很难通过实测建立目标的全极化 HRRP 数据库。本文利用 SolidWorks 3 维建模软件建立上述 7 类目标的 1:1 模型,然后导入 CST 电磁仿真软件进行仿真计算 4 种极化方式(HH, HV, VH, VV)下的全极化 HRRP, H, V 分别表示水平和垂直极化。对于对海雷达而言,上述 7 类目标在电磁仿真计算时都属于电极大目标,因此仿真效率较低。考虑到目标的对称性,同

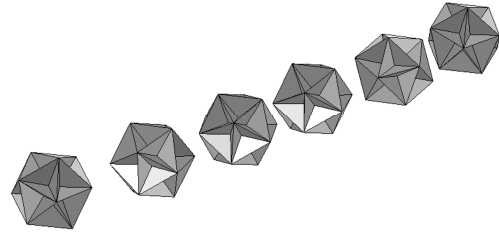


图 1 异型角反射体阵列结构示意图

表 1 各舰艇的主要结构参数

编号	舰长(m)	舰宽(m)	吃水深度(m)
舰艇 1	60.0	9.5	2.9
舰艇 2	135.0	16.0	4.5
舰艇 3	153.8	20.4	6.3
舰艇 4	157.0	19.0	6.0
舰艇 5	172.8	16.8	6.5
舰艇 6	182.8	24.1	8.1

时为兼顾仿真计算效率,本文选取舰首左舷一侧 $0^\circ \sim 30^\circ$, $75^\circ \sim 105^\circ$, $150^\circ \sim 180^\circ$ 这 3 个角域进行分析。

CST 仿真参数设置为:方位角为 $0^\circ \sim 30^\circ$, $75^\circ \sim 105^\circ$, $150^\circ \sim 180^\circ$ (舰首方向为 0° , 角度步长 0.1°),俯仰角为 90° ;中心频率为 10 GHz,带宽 100 MHz,频率采样点数为 128;采用软件默认的最优网格剖分尺寸,选择射线追踪算法进行求解。每类目标共仿真 903 个方位角,对应 903 组全极化 HRRP。方位角为 15° 时,舰艇 3 的仿真结果如图 2 所示。图中横坐标 d 表示距离,纵坐标 P 表示回波功率。

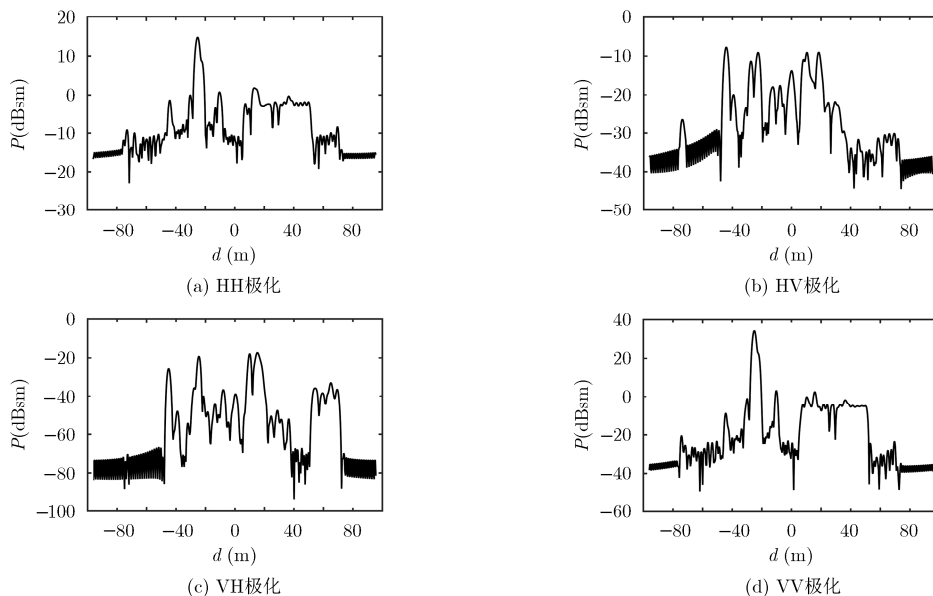


图 2 方位角 15° 时舰艇 3 的全极化 HRRP

可见不同极化方式下的 HRRP 具有较大差异, 尤其是同极化 HRRP 与交叉极化 HRRP 之间在波形和幅值上差异更大, 说明全极化 HRRP 中包含更全面的结构信息, 对全极化 HRRP 进行多特征提取可以提高对海雷达的目标识别性能。由于极化信息的处理过程都是建立在散射矩阵对称的基础上, 因此本文采用 Cameron 平均修正法对交叉极化通道的 HRRP 进行互易性修正^[6], 即取两者的平均值作为交叉极化通道的 HRRP, 修正后的全极化 HRRP 记为 HRRP_{HH} , HRRP_{HV} , HRRP_{VH} , HRRP_{VV} , 且有 $\text{HRRP}_{\text{HV}} = \text{HRRP}_{\text{VH}}$ 。

2.2 全极化 HRRP 特征提取

利用修正后的全极化 HRRP 提取以下 4 类特征: (1) 反映目标物理结构的平移不变特征, 共 21 个; (2) 反映目标 HRRP 散射随机性的熵特征, 共 4 个; (3) 反映目标与标准体散射矩阵的相似性特征, 共 7 个; (4) 反映目标与标准体散射能量的相似性特征, 共 7 个。

2.2.1 目标物理结构的平移不变特征 为了充分利用包含在 HRRP 中的目标精细结构信息, 从每个 HRRP 序列中提取 7 个与目标物理结构密相关的特征, 即径向长度 L , 散射重心 M , 强散射中心的数目 NP , 两个最强散射中心之间的距离 DPK , 最强散射中心距目标最前端的距离 DEP , 强散射中心的幅值分布熵 EA , 强散射中心的位置分布熵 EP 。其中, 前 5 个特征的提取可参见文献[3], 下面主要介绍特征 EA , EP 的提取。

将从 HRRP 序列中提取的 NP 个强散射中心, 按幅值递减顺序排列, 记为 $\mathbf{P} = [p_{m_1} p_{m_2} \cdots p_{m_{\text{NP}}}]$ 。其中, p_{m_i} 表示 HRRP 序列中第 i 强的散射中心的幅值, 下标 m_i 代表该散射中心对应的单元序号。令 $p'_{m_i} = p_{m_i} / \sum_{i=1}^{\text{NP}} p_{m_i}$, $m'_i = (m_i - n_1) / (n_2 - n_1)$, n_1, n_2 分别为 HRRP 序列中大于阈值的第 1 个和最后一个点所对应的序号, 则强散射中心的幅值分布熵 EA 可定义为

$$\text{EA} = -\sum_{i=1}^{\text{NP}} p'_{m_i} \log_2 p'_{m_i} \quad (1)$$

EA 反映了目标强散射中心幅值大小的离散程度, EA 越大说明幅值之间的差异越小; 反之差异越大。

强散射中心的位置分布熵 EP 定义为

$$\text{EP} = -\sum_{i=1}^{\text{NP}} m'_i \log_2 m'_i \quad (2)$$

EP 反映了目标强散射中心在径向尺度上位置分布的离散程度, EP 越大强散射中心在径向尺度上

均匀分布的可能性越大; 反之, 则说明强散射中心在径向尺度上的位置分布比较集中。

修正后每类目标在每个方位角对应 3 种不同收发极化的 HRRP, 每种极化方式对应 7 个平移不变特征, 因此, 每类目标在每个方位角总共对应 21 个反映目标物理结构的平移不变特征。

2.2.2 反映目标 HRRP 散射随机性的熵特征 2008 年, 文献[5]将用于全极化 SAR 分析中的散射熵 H 、平均散射角 α 和反熵 A 引入到基于全极化 HRRP 的目标识别领域。本文参考文献[6]中的方法, 提取反映目标 HRRP 散射随机性的 4 个熵特征, 即 H, α, A 和参数 P 。

H 值在 $[0, 1]$ 之间, 描述了目标散射的随机性, H 越大去极化程度越强; α 值在 $0^\circ \sim 90^\circ$ 之间, 反映了目标的主要散射机理, α 越大目标散射的各向异性越强; A 反映了次散射分量与最弱散射分量的能量比值; P 反映了主散射分量与次散射分量的能量比值。

2.2.3 反映目标与标准体散射矩阵及散射能量的相似性特征 舰船和角反射体等目标是由多个独立的子散射体所构成的复杂目标, 目标整体的散射特性是由各子散射体的相互作用共同决定, 研究目标与标准散射体的相似性对于目标间的彼此区分有一定价值。文献[8]从散射矩阵和散射能量两方面, 基于全极化 HRRP 定义了目标与 6 种标准体(即平板、二面角、水平偶极子、圆柱体、左螺旋体和右螺旋体)之间的相似性参数, 分别记为 r'_i 和 $r''_i, i=1, 2, \dots, 6$ 。考虑到舰船结构的特殊性, 本文引入倾斜 $\pi/4$ 的二面角(极化散射矩阵 $\mathbf{S}_7 = [0, 1; 1, 0]$), 按照文献[8]中方法提取目标与倾斜 $\pi/4$ 的二面角的散射矩阵及散射能量的相似性特征, 分别记为 r'_7 和 r''_7 。

综上所述, 每个目标在任意方位角可以提取 4 类总共 39 个特征, 本文仿真计算了 903 个方位角, 因此, 经特征提取后每个目标对应一个 39×903 的数据矩阵 $\mathbf{Z}_i, i=1, 2, \dots, 7$ 。虽然特征越多, 刻画目标越详实, 但并不是每个特征对分类识别都有贡献, 相反, 大量冗余特征会引发“维数灾难”且会增加拟合的风险, 因此有必要进行特征选择, 只保留对分类识别贡献较大的特征。

3 全极化 HRRP 特征选择

基于 MI 的特征选择算法利用互信息来定量分析特征间或特征与类标签向量间的线性或非线性的关系, 离散变量 X 与 Y 的互信息可以表示为

$$I(X; Y) = \sum_{y \in \Omega_Y} \sum_{x \in \Omega_X} p(x, y) \log_2 \{p(x, y) / [p(x)p(y)]\} \quad (3)$$

式中, Ω_X 和 Ω_Y 为 X 和 Y 的样本空间; $p(x)$, $p(y)$, $p(x,y)$ 分别为 X , Y 以及 (X,Y) 的概率密度函数。

MI 也可以用熵来表示, 即

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X,Y) \end{aligned} \quad (4)$$

式中, $H(X)$ ($H(Y)$), $H(X|Y)$ 分别为熵和条件熵。

$$\left. \begin{aligned} H(X) &= - \sum_{x \in \Omega_X} p(x) \log_2 p(x) \\ H(X|Y) &= - \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} p(x,y) \log_2 p(x|y) \\ H(X,Y) &= - \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} p(x,y) \log_2 p(x,y) \end{aligned} \right\} \quad (5)$$

由式(4)和式(5)可知互信息还可以理解为: 已知变量 Y 的情况下, 变量 X 的不确定度的减少量。

3.1 最小冗余最大相关特征选择

分析文献[12]中列举的 17 种基于 MI 的特征选择算法, 其中绝大多数算法都是围绕冗余度和相关性展开研究, 最为成功的则是最小冗余最大相关算法 (Minimum Redundancy Maximum Relevance, MRMR)。特征集 $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ 由 n 个特征组成, C 表示类标签变量, 特征 X_i 与 C 的相关性用 MI 表示为

$$\text{Rel}(X_i) \triangleq I(X_i; C) \quad (6)$$

特征 X_i 相对已选择特征集 \mathbf{S}_{i-1} 的冗余度用 MI 表示为

$$\text{Red}(X_i | \mathbf{S}_{i-1}) \triangleq \sum_{X_j \in \mathbf{S}_{i-1}} I(X_i; X_j) / |\mathbf{S}_{i-1}| \quad (7)$$

式中, $|\mathbf{S}_{i-1}|$ 代表 \mathbf{S}_{i-1} 中的特征数。

MRMR 是一次只选择一个特征的贪婪算法, 第 i 次选择的特征需要满足式(8)的目标函数最大化要求, 即

$$\max_{X_i \in \mathbf{X} \setminus \mathbf{S}_{i-1}} \{ \text{Rel}(X_i) - \text{Red}(X_i | \mathbf{S}_{i-1}) \} \quad (8)$$

广义 MRMR 算法可以归纳为

$$\text{MRMR: } \max_{X_i \in \mathbf{X} \setminus \mathbf{S}_{i-1}} \left\{ I(X_i; C) - \alpha \sum_{X_j \in \mathbf{S}_{i-1}} I(X_i; X_j) \right\} \quad (9)$$

式中, α 为用于平衡相关项和冗余项的加权因子。

由式(9)可知, α 的选取是决定所选特征能否在最小冗余和最大相关两方面达成良好折中的关键。 α 过大, 右侧冗余项将发挥主导作用, 算法倾向于选择具有最小冗余度的特征; 反之, 如果 α 过小, 算法倾向于选择具有最大分类相关性的特征。针对 α 选取的研究很多, 比较有代表性的当属以下 4 种算法:

MIFS 和 MIFS-U: 通过试验人为选择 α [14]。

mRMR [15]: $\alpha = 1 / |\mathbf{S}_{i-1}|$ 。

NMIFS-1 [16]: $\alpha(X_i; X_j) = 1 / \{ |\mathbf{S}_{i-1}| \times \min(H(X_i), H(X_j)) \}$, $X_j \in \mathbf{S}_{i-1}$ 。

目前, 基于 MI 的特征选择算法在涉及 α 选取时, 或需进行大量实验, 或无法从根本上解决相关项和冗余项的平衡问题。另外, 据笔者所知目前基于 MI 的特征选择算法都具有递增性, 即每次只能选择一个特征, 且已选择的特征在后续过程中不能被删除, 从而决定了所得特征子集的次优性。针对上述两个问题, 本文提出一种基于归一化互信息 (NMI) 并利用模拟退火 (SA) 算法进行优化的全局最优特征选择算法, 记为 NMI-SA。

3.2 NMI-SA 算法

NMI-SA 算法同样以最小冗余最大相关准则为出发点, 与传统 MRMR 算法不同的是 NMI-SA 可以评价任意特征子集的优良程度, 即利用多个特征与类标签向量的归一化联合互信息来衡量特征子集的相关性, 并利用特征间的平均归一化互信息来衡量特征子集的冗余度, 最后利用 SA 算法搜索全局最优的特征子集。

3.2.1 基于归一化互信息的适应度 为了寻找合适的互信息归一化方法, 首先需要分析互信息的上界, 由于连续变量可以量化为离散变量, 因此选取两个离散随机变量 X 和 Y 进行分析。由式(4)可知

$$I(X;Y) \leq \min\{H(X), H(Y)\} \quad (10)$$

将詹森不等式用于式(5)中的熵定义可得

$$H(X) \leq \log_2 \left\{ \sum_{x \in \Omega_X} p(x) \frac{1}{p(x)} \right\} = \log_2 |\Omega_X| \quad (11)$$

综合式(10)和式(11)可得

$$I(X;Y) \leq \min\{\log_2 |\Omega_X|, \log_2 |\Omega_Y|\} \quad (12)$$

为了保证每个特征的取值个数相同, 用同一量化水平 N 对所有特征进行量化处理, 量化过程中逐一增加量化位数 N , 直至所有样本的最大量化误差小于预设的期望量化误差 ε 。本文取 $\varepsilon = 0.01$, 因为实验发现更小的 ε 对提高分类精度几乎没有贡献, 反而会带来额外的计算负担。量化处理后, 对任意特征 $|\Omega_X| = N$, 因此 $I(X;Y) \leq \log_2 N$, 显然 $\log_2 N$ 是与特征无关的互信息上界。进而将特征之间的归一化互信息定义为

$$NI(X;Y) = I(X;Y) / \log_2 N \quad (13)$$

显然, 式(13)的取值范围为 $[0,1]$ 。

设特征子集 $\mathbf{S} = \{X_{(1)}, X_{(2)}, \dots, X_{(m)}\}$ 包含 m 个特征, \mathbf{S} 与类标签变量 C 的相关程度可用联合互信息表示, 即

$$\begin{aligned}
 \text{Rel}(\mathbf{S}) &\triangleq I(C; \mathbf{S}) = I(C; X_{(1)}, X_{(2)}, \dots, X_{(m)}) \\
 &= H(C) - H(X | X_{(1)}, X_{(2)}, \dots, X_{(m)}) \\
 &= H(C) + H(X_{(1)}, X_{(2)}, \dots, X_{(m)}) \\
 &\quad - H(C, X_{(1)}, X_{(2)}, \dots, X_{(m)}) \quad (14)
 \end{aligned}$$

显然, $I(C; \mathbf{S}) \leq H(C)$, 因此, 可以用 $\log_2 |\Omega_C|$ 进行归一化, 记为归一化联合互信息 $NJI(C; \mathbf{S})$, 即

$$\text{Rel}(\mathbf{S}) \triangleq NJI(C; \mathbf{S}) = I(C; \mathbf{S}) / \log_2 |\Omega_C| \quad (15)$$

另外, 特征子集 \mathbf{S} 的冗余度可用特征间的平均归一化互信息来表征, 即

$$\text{Red}(\mathbf{S}) \triangleq \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m NI(X_{(i)}; X_{(j)}) \quad (16)$$

由式(15)和式(16)可知, $\text{Rel}(\mathbf{S})$ 与 $\text{Red}(\mathbf{S})$ 取值范围均为 $[0, 1]$, 可定义如下特征子集适应度函数:

$$f(\mathbf{S}) = \text{Rel}(\mathbf{S}) - \text{Red}(\mathbf{S}) \quad (17)$$

利用模拟退火算法进行全局寻优等价于搜索使 $f(\mathbf{S})$ 取最大值的特征子集 \mathbf{S} 。

为了便于后续对比研究, 在此基于归一化互信息定义如下特征选择算法, 即

$$\text{NMIFS-2: } \max_{X_i \in \mathbf{X} \setminus \mathbf{S}_{i-1}} \left\{ NI(X_i; C) - \sum_{X_j \in \mathbf{S}_{i-1}} NI(X_i; X_j) \right\} \quad (18)$$

式中, $NI(X_i; C) = I(X_i; Y) / \log_2 |\Omega_C|$ 。

3.2.2 NMI-SA 算法实现 SA 算法包括初温设定、Metropoils 抽样和控制参数的下降 3 个过程。能量就是代价函数, 要得到的最优解就是能量最低态, 即使代价函数取最小值的解。Metropoils 准则以一定的概率接受恶化解, 从而避免陷入局部最优解^[17]。取代价函数 $F(\mathbf{S}) = -f(\mathbf{S})$, NMI-SA 的伪代码如表 2 所示, 本文取 $T_0 = 100$, $T_{\min} = 0.1$ 。

表 2 NMI-SA 的伪代码

<p>输入: 学习数据集 \mathbf{D}; 特征集 $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$; 特征子集的特征数 m; 当前温度 T; 初始温度 T_0; 最小温度 T_{\min}。</p> <p>初始化阶段: 令 $T = T_0$, 任取初始解 $\mathbf{S}_1 = \{X_{(1)}, X_{(2)}, \dots, X_{(m)}\}$, $\mathbf{S}_2 = \mathbf{X} - \{X_{(1)}, X_{(2)}, \dots, X_{(m)}\}$。</p> <p>特征子集选择阶段: 重复步骤(1)~步骤(5), 直到 $T < T_{\min}$。</p> <p>(1)从 \mathbf{S}_2 中随机选取一个特征替换 \mathbf{S}_1 中的一个特征, 记为 \mathbf{S}_1^*;</p> <p>(2)如果 $F(\mathbf{S}_1^*)$ 小于当前最优解对应的代价函数值, 将 \mathbf{S}_1^* 存为最优解;</p> <p>(3)计算 \mathbf{S}_1^* 的增量 $dF = F(\mathbf{S}_1^*) - F(\mathbf{S}_1)$;</p> <p>(4)若 $dF < 0$, 则接受 \mathbf{S}_1^* 作为新的当前解, 即 $\mathbf{S}_1 = \mathbf{S}_1^*$, 否则计算 \mathbf{S}_1^* 的接受概率 $P = \exp(-dF/T)$;</p> <p>若 $P > \text{rand}$, 令 $\mathbf{S}_1 = \mathbf{S}_1^*$, 否则保留当前解 \mathbf{S}_1。(注: rand 为 $[0, 1]$ 之间的一个随机数);</p> <p>(5)更新 $\mathbf{S}_2 = \mathbf{X} - \mathbf{S}_1$; 降温 $T = 0.95T$。</p> <p>输出: 最优特征子集。</p>
--

4 实验

实验环境为 Core i5 处理器、主频 3.3 GHz、内存 4 GB, 选用 MATLAB R2016a, 选择 PRtools 5.0 中的 k -近邻分类器(k -nn, k 取 3)进行分类识别。分类精度用 5×2 交叉校验的平均值来表征, 即重复进行 5 次 2 重交叉校验实验, 每次实验时将样本集随机分为等大的两部分, 轮流充当训练集和测试集。首先进行实验对比 NMI-SA 与 mRMR, NMIFS-1, NMIFS-2 的特征选择性能; 然后, 利用性能最优的特征选择算法重点分析全极化 HRRP 的 39 个特征, 旨在寻找一个使适应度 $f(\mathbf{S})$ 和分类正确率尽可能高的特征子集。

4.1 性能对比实验

设计实验对比 NMI-SA 与 mRMR, NMIFS-1, NMIFS-2 的特征选择性能, 为了保证试验的客观性, 除了以目标全极化 HRRP 特征集作为实验数据集外, 还从 UCI 公共数据库中选取了 9 个数据集, 分别为 Hill Valley, Ionosphere, Isolet, Libras Movement, Landsat Satellite, Letter, Sonar, Spambase, Vehicle。上述数据集的样本数、类别数、特征数或特征类型等信息参见 <http://archive.ics.cui.edu/ml/datasets.html>。

4 种特征选择算法在不同数据集上选择不同尺寸的特征子集, 利用 k -nn 进行分类后的正确率以及各算法的排名如表 3 所示(平局时用取名次的均值, 例如出现并列第 3 时, 排名记为 3.5)。另外, 在数据集名称旁边还给出了利用全部特征进行分类时的正确率作为参考。表 3 中通过加黑突出显示性能最优的算法。

4 种算法在各数据集上的平均排名如图 3 所示, 可见 NMI-SA 的排名比较靠前, 整体上优于其余 3 种算法。为得到更具统计意义的结论进行 Friedman 检验。Friedman 检验对比各种算法的平均排名, 即

$$R_j = \frac{1}{N} \sum_i r_i^j, \quad j = 1, 2, \dots, 4 \quad (19)$$

式中, r_i^j 为第 j 种算法在第 i 个数据集上的排名; $N = 10$ 为数据集总数。

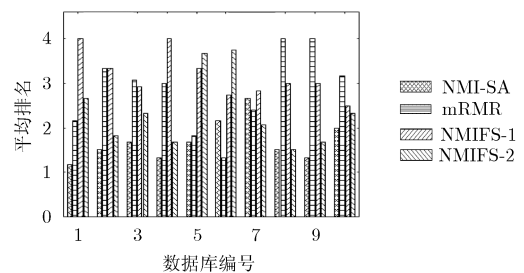


图 3 算法在各数据集上的平均排名

表3 各特征选择算法的分类正确率及排名表

全极化HRRP(90.13)					Hill Valley(50.70)				
特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2	特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2
5	80.30/1	77.10/3	69.39/4	77.62/2	5	52.46/1	49.64/3.5	49.64/3.5	51.70/2
10	87.83/1	87.61/2	77.84/4	86.55/3	10	53.14/1	50.89/3.5	50.89/3.5	52.44/2
15	90.09/2	91.10/1	79.77/4	87.10/3	15	52.52/1	50.87/3.5	50.87/3.5	51.06/2
20	91.35/1	90.82/2	81.23/4	87.19/3	20	53.47/4	53.56/2.5	53.56/2.5	53.60/1
25	93.19/1	92.84/2	81.33/4	87.19/3	25	53.25/1	52.36/3.5	52.36/3.5	52.74/2
30	91.62/1	91.05/3	87.92/4	91.31/2	30	54.01/1	53.56/3.5	53.56/3.5	53.84/2
均值/排名	1.17	2.17	4	2.67	平均排名	1.5	3.33	3.33	1.83
Ionosphere(89.29)					Isolet(85.58)				
特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2	特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2
5	89.77/1	87.98/4	88.95/3	89.34/2	5	42.19/2	31.67/3	29.78/4	44.63/1
10	89.71/1	89.57/2.5	89.57/2.5	88.66/4	10	55.02/2	49.95/3	45.16/4	57.5/1
15	89.20/2	88.26/4	89.12/3	89.51/1	15	64.32/1	53.07/3	48.04/4	64.10/2
20	89.14/3	89.46/2	89.57/1	88.38/4	20	68.24/1	58.19/3	50.63/4	66.56/2
25	89.21/2	89.17/3	89.06/4	89.23/1	25	68.79/1	61.04/3	53.87/4	68.30/2
30	90.29/1	89.18/3	88.83/4	89.35/2	30	71.03/1	62.45/3	55.08/4	70.81/2
平均排名	1.67	3.08	2.92	2.33	平均排名	1.33	3	4	1.67
Libras Movement(70.56)					Landsat Satellite(90.14)				
特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2	特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2
5	51.11/2	52.67/1	41.94/3	38.72/4	5	84.61/4	85.15/1	85.13/2	84.74/3
10	62.78/1	62.39/2	51.72/3	47.94/4	10	87.84/3	87.91/1	87.87/2	87.09/4
15	64.33/1	63.11/2	55.78/3	52.17/4	15	88.59/2	88.73/1	88.54/3	88.20/4
20	65.56/1	64.50/2	56.89/3	56.39/4	20	89.43/1	89.22/2	89.07/3	88.82/4
25	67.83/1	63.83/2	60.11/4	60.78/3	25	89.83/1	89.76/2	89.61/3.5	89.61/3.5
30	68.00/1	65.61/2	61.94/4	64.00/3	30	90.12/2	90.15/1	89.99/3	89.82/4
平均排名	1.17	1.83	3.33	3.67	平均排名	2.17	1.33	2.75	3.75
Letter(91.46)					Sonar(78.85)				
特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2	特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2
2	14.94/4	15.16/2	15.16/2	15.16/2	5	65.37/2	61.84/4	65.04/3	71.70/1
4	59.95/1	57.43/3	53.63/4	58.13/2	10	70.86/2	63.56/4	70.06/3	77.77/1
6	79.66/3	80.29/2	74.85/4	80.34/1	15	74.29/2	68.73/4	73.20/3	78.02/1
8	86.48/4	88.65/2.5	88.75/1	88.65/2.5	20	81.90/1	70.65/4	77.54/3	80.02/2
10	92.48/3	92.76/2	92.32/4	92.86/1	25	81.26/1	76.54/4	77.12/3	78.94/2
12	92.81/1	92.76/3	92.79/2	92.75/4	30	78.64/1	75.76/4	76.15/3	77.20/2
平均排名	2.67	2.42	2.83	2.08	平均排名	1.5	4	3	1.5
Spambase(88.89)					Vehicle(66.81)				
特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2	特征数	NMI-SA	mRMR	NMIFS-1	NMIFS-2
5	78.74/2	59.29/4	70.97/3	81.26/1	2	48.82/4	61.16/1	56.16/2	53.14/3
10	85.43/1	71.10/4	79.15/3	83.44/2	4	56.40/1	55.93/2	53.64/4	55.51/3
15	88.40/1	75.32/4	81.60/3	88.31/2	6	62.74/1	60.61/4	61.54/3	61.91/2
20	88.43/2	81.39/4	84.72/3	88.59/1	8	64.93/2	61.63/4	65.16/1	63.86/3
25	89.66/1	84.47/4	86.15/3	88.78/2	10	65.64/3	63.27/4	65.80/2	66.26/1
30	90.04/1	86.23/4	87.44/3	89.14/2	12	69.81/1	64.11/4	67.24/3	69.31/2
平均排名	1.33	4	3	1.67	平均排名	2	3.17	2.5	2.33

经检验得 $p = 0.003 < 0.01$ ，说明 4 种算法存在显著差异。为了找到具体差异需要进行 pot-hoc 检验，本文主要关注 NMI-SA 是否显著优于其余 3 种算法，因此，采用 Bonferroni-Dunn 检验——即两种算法具有显著差异当两者之间的平均排名差大于临界值 CD：

$$CD = q_\alpha \sqrt{k(k+1)/6N} \quad (20)$$

式中， q_α 可通过查询 “The Studentized Range Statistic” 表得到； $k = 4$ 为所要验证的算法数。

取显著性水平 $\alpha = 0.1$ ， $q_\alpha = 1.834$ ，代入式(20)得差异临界值 CD 为 1.059，将 NMI-SA 与另外 3 种算法进行 Bonferroni-Dunn 检验的可视化结果如图 4 所示，可见 NMI-SA 算法的平均排名是最靠前的，且 NMI-SA 算法显著优于位于 CD 区间外的 mRMR 和 NMIFS-1 算法，但与 NMIFS-2 相比优势并不显著。

4.2 全极化 HRRP 特征选择

4.1 节验证了 NMI-SA 算法整体优于另外 3 种算法，本节利用 NMI-SA 算法从全极化 HRRP 的 39 个特征中挑选出辨别力强、冗余度低的特征子集。本文仿真建立了方位角为 $0^\circ \sim 30^\circ$ ， $75^\circ \sim 105^\circ$ ， $150^\circ \sim 180^\circ$ 这 3 个角域的 HRRP 全极化特征数据库，在特征选择时分别对这 3 个角域以及 3 个角域构成的整体进行研究。特征子集的尺寸由 1 增加到 39(步长为 1)，利用 NMI-SA 算法选择出不同尺寸下的最优特征子集，利用 k -nn 分类器输出的不同特征集尺寸下的分类正确率如图 5 所示。

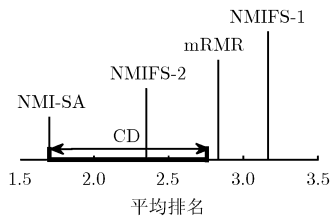


图 4 Bonferroni-Dunn 检验结果

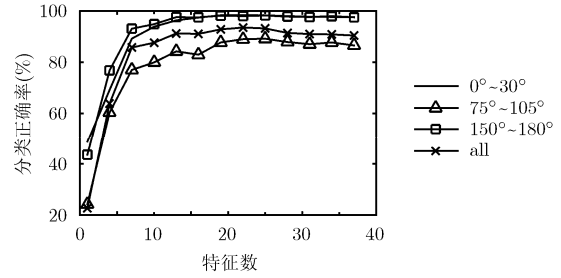


图 5 不同特征集尺寸下的分类正确率

由图 5 可见，对所有角域当特征子集的特征数达到 25 之后，分类正确率不再随特征数的增加而提高，甚至还会出现下降，从而确定最优特征子集的尺寸为 25。这一现象验证了特征并不是越多越好，必须通过特征选择去除冗余特征，提高特征集的泛化能力。另外，考虑到 HRRP 具有方位敏感性，很难保证提取的所有特征对方位均不敏感，因此有必要分别研究不同 HRRP 角域支撑区下的最优特征子集，基于不同角域支撑区利用 NMI-SA 算法确定的最优特征子集如表 4 所示，表中标记 a~d 分别表示基于角域 $0^\circ \sim 30^\circ$ ， $75^\circ \sim 105^\circ$ ， $150^\circ \sim 180^\circ$ 和上述所有角域所选中的特征。

由表 4 可知，基于不同角域所选的最优特征子集并不相同，尤其是 $75^\circ \sim 105^\circ$ 对应的最优特征子集与 $0^\circ \sim 30^\circ$ ， $150^\circ \sim 180^\circ$ 对应的特征子集差别较大(分别有 8 个、10 个不同特征)，而 $0^\circ \sim 30^\circ$ 与 $150^\circ \sim 180^\circ$ 对应的最优特征子集差别较小(只有 4 个不同特征)。这说明不同角域支撑区对应不同的最优特征子集，由于实际应用中无法确定雷达与海上目标的相对态势，为了尽可能松弛方位敏感性，应选取尽可能大的角域进行特征子集的优选。为此，本文以 3 个角域构成的全角域对应的最优特征子集作为最终结果，该子集中 4 类特征分别占 16 个，3 个，4 个和 2 个，说明第 1 类特征辨别力最强；另外，在第 3 和第 4 类特征中目标与平板、二面角、圆柱体和倾斜 $\pi/4$ 的二面角的结构或散射能量相似性参数作用比较突出。

表 4 基于不同角域确定的最优特征子集表

特征 1-8	特征 9-16	特征 17-24	特征 25-32	特征 33-39					
L(HH)	a,c,d	NP(VV)	a,b,c,d	EA(HV)	b,d	P	c	r_1''	a,b,c,d
L(HV)	b	DPK(HH)	c,d	EA(VV)	a,b,c,d	r_1'	a,b,c,d	r_2''	a,c
L(VV)	a,c,d	DPK(HV)	b	EP(HH)	a,c	r_2'	a,b,c,d	r_3''	b
M(HH)	a,c,d	DPK(VV)	a,c	EP(HV)	a,b,d	r_3'	b	r_4''	a,b,d
M(HV)	a,b,c,d	DEP(HH)	a,b,c,d	EP(VV)	b,c,d	r_4'	a,b,c,d	r_5''	
M(VV)	a,c,d	DEP(HV)	a,b,d	H	a,b,c,d	r_5'		r_6''	
NP(HH)	a,b,c,d	DEP(VV)	b,c,d	α	a,b,c,d	r_6'		r_7''	a,b
NP(HV)	a,b,c,d	EA(HH)	a,c	A	b,d	r_7'	a,b,c,d		

5 结束语

利用 CST 软件仿真建立了 7 类海上目标的在不同方位角下的全极化 HRRP 数据库；在此基础上，提取了 4 类共 39 个特征。为了从根本上解决现有基于互信息的特征选择算法的次优性以及平衡因子选择方面的难题，本文提出一种基于归一化互信息并利用模拟退火算法进行优化的全局最优特征选择算法—NMI-SA。基于 HRRP 数据集以及 9 个 UCI 数据集，并利用 k -近邻分类器进行对比实验，验证了 NMI-SA 算法的可行性和优越性。利用 NMI-SA 算法对全极化 HRRP 的 39 个特征进行重点分析，根据分类识别率的变化规律确定特征子集的最优尺寸为 25；分析了 NMI-SA 算法的方位敏感性，为了尽可能松弛方位敏感性，基于 3 个角域构成的全角域优选出一个尺寸为 25 的最优特征子集。本文只研究了模拟退火与归一化互信息的结合，后续将在此基础上研究使用不同搜索算法对特征选择性能的影响。

参考文献

- [1] 冯博, 陈渤, 王鹏辉, 等. 利用稳健字典学习的雷达高分辨距离像目标识别算法[J]. 电子与信息学报, 2015, 37(6): 1457-1462. doi: 10.11999/JEIT141227.
FENG Bo, CHEN Bo, WANG Penghui, et al. Radar high resolution range profile target recognition algorithm via stable dictionary learning[J]. *Journal of Electronics & Information Technology*, 2015, 37(6): 1457-1462. doi: 10.11999/JEIT141227.
- [2] 郭尊华, 李达, 张伯彦. 雷达高距离分辨率一维像目标识别[J]. 系统工程与电子技术, 2013, 35(1): 53-60. doi: 10.3969/j.issn.1001-506X.2013.01.09.
GUO Zunhua, LI Da, and ZHANG Boyan. Survey of radar target recognition using one-dimensional high range resolution profiles[J]. *Systems Engineering and Electronics*, 2013, 35(1): 53-60. doi: 10.3969/j.issn.1001-506X.2013.01.09.
- [3] PICHER C and KHOTANZAD A. Nonlinear classifier combination for a maritime target recognition task[C]. Proceedings of the IEEE Radar Conference, Pasadena, 2009: 873-877. doi: 10.1109/RADAR.2009.4976923.
- [4] 刘盛启, 占荣辉, 翟庆林, 等. 基于联合稀疏性的多视全极化 HRRP 目标识别方法[J]. 电子与信息学报, 2016, 38(7): 1724-1730. doi: 10.11999/JEIT151019.
LIU Shengqi, ZHAN Ronghui, ZHAI Qinglin, et al. Multi-view polarization HRRP target recognition based on joint sparsity[J]. *Journal of Electronics & Information Technology*, 2016, 38(7): 1724-1730. doi: 10.11999/JEIT151019.
- [5] BERIZZI F, MARTORELLA M, CAPRIA A, et al. H/ α polarimetric features for man-made target classification[C]. Proceedings of the IEEE Radar Conference, Rome, 2008: 1-6. doi: 10.1109/RADAR.2008.4721003.
- [6] 杨磊, 王晓丹, 张玉玺, 等. 基于多极化特征提取和 SVM 的目标识别方法[J]. 现代防御技术, 2012, 40(5): 150-155. doi: 10.3969/j.issn.1009-086x.2012.05.029.
YANG Lei, Wang Xiaodan, ZHANG Yuxi, et al. Radar target recognition approach based on multi polarization multi target feature extraction and SVM[J]. *Modern Defence Technology*, 2012, 40(5): 150-155. doi: 10.3969/j.issn.1009-086x.2012.05.029.
- [7] 雷蕾, 王晓丹, 邢雅琼, 等. 结合 SVM 和 DS 证据理论的多极化 HRRP 分类研究[J]. 控制与决策, 2013, 28(6): 861-866. doi: 10.13195/j.cd.2013.06.63.leil.011.
LEI Lei, WANG Xiaodan, XING Yaqiong, et al. Multi-polarized HRRP classification by SVM and DS evidence theory[J]. *Control and Decision*, 2013, 28(6): 861-866. doi: 10.13195/j.cd.2013.06.63.leil.011.
- [8] 郭雷. 宽带雷达目标极化特征提取与核方法识别研究[D]. [博士学位论文], 国防科学技术大学, 2009: 15-49.
GUO Lei. Wideband radar target polarimetric feature extraction and recognition method based on kernel method [D]. [Ph.D. dissertation], National University of Defense Technology, 2009: 15-49.
- [9] LIU H, SUN J, LIU L, et al. Feature selection with dynamic mutual information[J]. *Pattern Recognition*, 2009, 42(7): 1330-1339. doi: 10.1016/j.patcog.2008.10.028.
- [10] UNLER A, MURAT A, and CHINNAM R B. mr 2 PSO : a maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification[J]. *Information Sciences*, 2011, 181(20): 4625-4641. doi: 10.1016/j.ins.2010.05.037.
- [11] GARCIA M, GOMEZ F, MELIAN B, et al. High-dimensional feature selection via feature grouping: a variable neighborhood search approach[J]. *Information Sciences*, 2016, 326(C): 102-118. doi: 10.1016/j.ins.2015.07.041.
- [12] BROWN G, POCOCK A, ZHAO M J, et al. Conditional likelihood maximization: a unifying framework for information theoretic feature selection[J]. *Journal of Machine Learning Research*, 2012, 13(1): 27-66.
- [13] LYSIK R, KURZYNSKI M, and WOLOSZYNSKI T. Optimal selection of ensemble classifiers using measures of competence and diversity of base classifiers[J]. *Neurocomputing*, 2014, 126(1): 29-35. doi: 10.1016/j.neucom.2013.01.052.
- [14] KWAK N and CHOI C H. Input feature selection for classification problems[J]. *IEEE Transactions on Neural Networks*, 2002, 13(1): 143-159. doi:10.1109/72.977291.
- [15] PENG H, LONG F, and DING C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2005, 27(8): 1226-1238. doi: 10.1109/TPAMI.2005.159.
- [16] ESTEVEZ P A, TESMER M, PEREZ C A, et al. Normalized mutual information feature selection[J]. *IEEE Transactions on Neural Networks*, 2009, 20(2): 189-201. doi: 10.1109/TNN.2008.2005601.
- [17] ISAKOV S V, ZINTCHENKO I N, RONNOW T F, et al. Optimised simulated annealing for icing spin glasses[J]. *Computer Physics Communications*, 2015, 192: 265-271. doi: 10.1016/j.cpc.2015.02.015.

范学满: 男, 1989 年生, 博士生, 研究方向为精确制导与对抗。
胡生亮: 男, 1974 年生, 教授, 研究方向为无源对抗。
贺静波: 男, 1979 年生, 讲师, 研究方向为随机微分理论及应用。