

引用格式:王晨阳. PCA和GA-BP结合的地磁导航适配区选择方法[J]. 电光与控制, 2018, 25(6):110-114. WANG C Y. Geomagnetic navigation matching area selection based on PCA and GA-BP neural network[J]. Electronics Optics & Control, 2018, 25(6):110-114.

## PCA和GA-BP结合的地磁导航适配区选择方法

王晨阳

(南京理工大学自动化学院, 南京 210094)

**摘要:** 由于地磁图适配区的选择是影响地磁导航定位精度的重要因素, 因此提出一种基于主成分分析法(PCA)和GA-BP神经网络相结合的地磁背景场适配/非适配区自动识别和分类的方法。首先利用PCA对地磁特征参数进行分析, 选择出独立的、并且包含主成分的特征参量, 其次构建GA-BP神经网络模型, 建立地磁特征参数和匹配性能的对应关系, 从而实现适配/非适配区的划分。通过多次仿真试验, 证明了采用该方法能够选择出较好的适配区域, 提高地磁导航定位精度。

**关键词:** 地磁导航; 特征参数; 主成分分析; GA-BP神经网络; 适配区选择

**中图分类号:** TP391.41 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.06.024

## Geomagnetic Navigation Matching Area Selection Based on PCA and GA-BP Neural Network

WANG Chen-yang

(School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)

**Abstract:** The selection of suitable matching area of geomagnetic map is important for ensuring the positioning accuracy of geomagnetic navigation. This paper puts forward a method for the automatic recognition and classification of the suitable and unsuitable matching areas of geomagnetic background field based on Principal Component Analysis (PCA) and GA-BP neural network. To select independent characteristic parameters containing the main components, PCA is used to analyze the geomagnetic characteristic parameters. Then, the GA-BP neural network model is constructed, and the correspondence between the geomagnetic characteristic parameters and matching performance is established, so as to realize the recognition and classification of suitable and unsuitable matching areas. Simulation results show that this method can efficiently find out a more effective matching area, and improve the positioning accuracy of geomagnetic navigation.

**Key words:** geomagnetic navigation; characteristic parameter; principal component analysis; GA-BP neural network; matching area selection

### 0 引言

近年来,地磁导航以其全地域、全天候、无辐射等特征,在舰船、无人机、航天器等系统得到广泛应用,优势日渐凸显<sup>[1]</sup>。地磁导航的核心在于获取精准的匹配位置信息,匹配算法并不是唯一影响匹配位置信息的因素,地磁区域的适配性也是影响匹配精度的另一个重要因素<sup>[2]</sup>。可以通过分析地磁图适配性,将信息丰富、地磁特性明显的区域选为地磁适配区,进而达到增强导航准确性的目的<sup>[3]</sup>。

目前,国内外关于适配性问题的研究主要集中于定性和定量这两个方面。文献[4]以图像中的每个像元为考察对象,综合运用了线性加权求和与属性相乘的方法来构建像元的适配性度量准则;文献[5]利用主成分分析法对适配性问题进行了深入的研究;文献[6]利用公式判定策略将分类准则抽象成数学公式,当候选适配区满足公式成立的条件时,即可被判定为适配区;文献[7]利用支持向量机(Support Vector Machine, SVM)实现了对地磁场区域的分类。

为了进一步提高适配区域选择的精度和可拓展性,本文从特征选择和参数优化两个角度出发,将PCA和GA-BP神经网络相结合,对地磁图进行适配/非适配区的划分,以选择出更加利于匹配的适配区域。

收稿日期:2017-06-05

修回日期:2017-07-25

作者简介:王晨阳(1993—),女,陕西渭南人,硕士生,研究方向为地磁导航相关技术。

## 1 地磁特征参数的选取

### 1.1 地磁图特征

从已有研究来看,表征地磁图特征的参数较多,如地磁粗糙度、标准差、地磁信息熵等,均从不同的角度表征了地磁图的特征。由于单个的特征参量表征地磁图的适配性有较大的缺陷,故应综合考虑多特征参数来衡量地磁适配性<sup>[8]</sup>。本文从适配性能分析的角度出发,从宏观起伏特征和微观破碎特征中,各选取3种常用的具有代表性的特征参数,与自相似特征中的相关系数共7个特征一同作为研究对象。

设候选匹配区的格网大小为 $M \times N$ ( $M$ 为纬度跨度, $N$ 为经度跨度), $f(i,j)$ 为点 $(i,j)$ 处的地磁场强度值。

#### 1) 地磁费歇信息量 $F_{ic}$ 。

地磁场包含的地磁信息量可以用地磁费歇信息量来度量。定义为

$$F_{ic} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \|\nabla h(x_{i,j})\|^2} \quad (1)$$

式中: $x_{i,j}$ 为平面位置点坐标,平均分布在地磁场区域内; $\nabla h(x_{i,j})$ 为第 $x_{i,j}$ 点的磁场梯度。地磁费歇信息量可取的值越大,则该区域有越丰富的地磁信息。

#### 2) 地磁标准差 $\sigma$ 。

地磁标准差反映了整个区域的起伏程度和该区域地磁场分布的离散程度,值越大,越有利。定义为<sup>[9]</sup>

$$\sigma = \sqrt{\frac{1}{MN-1} \sum_{i=1}^M \sum_{j=1}^N (f(i,j) - \bar{f})^2} \quad (2)$$

式中, $\bar{f} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N f(i,j)$ 为该区域地磁场的均值。

#### 3) 地磁累加梯度 $T$ 。

地磁累加梯度能将地磁场区域的差异性和复杂性有效地反映出来,累积梯度越大,越有利。表示为<sup>[9]</sup>

$$T = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \sqrt{T_x(i,j)^2 + T_y(i,j)^2} \quad (3)$$

式中, $T_x(i,j)$ 和 $T_y(i,j)$ 分别表示点 $(i,j)$ 处纬度和经度方向的梯度值。

#### 4) 地磁信息熵 $H$ 。

地磁信息熵用来度量地磁场的起伏特征和地磁信息的丰富程度,值越小,特征越明显。熵的单位为bit,定义为<sup>[10]</sup>

$$H = - \sum_{i=1}^M \sum_{j=1}^N p(i,j) \cdot \ln p(i,j) \quad (4)$$

式中, $p(i,j) = \frac{f(i,j)}{\sum_{i=1}^M \sum_{j=1}^N f(i,j)}$ 。

#### 5) 地磁粗糙度方差 $R$ 。

地磁粗糙度方差表征区域地磁场的局部起伏和平

均光滑程度,值越大越有利。定义为<sup>[9]</sup>

$$R = r/\sigma \quad (5)$$

式中, $r$ 为粗糙度, $r = (r_x + r_y)/2$ , $r_x = \left( \frac{1}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^{N-1} [f(i,j) - f(i,j+1)]^2 \right)^{1/2}$ 和 $r_y = \left( \frac{1}{(M-1)N} \sum_{i=1}^{M-1} \sum_{j=1}^N [f(i,j) - f(i+1,j)]^2 \right)^{1/2}$

分别表示该区域 $x$ 和 $y$ 方向的粗糙度。

#### 6) 地磁坡度方差 $S$ 。

地磁坡度方差衡量整个区域地磁坡度情况。定义为<sup>[9]</sup>

$$S = \sqrt{\frac{1}{(M-2)(N-2)-1} \sum_{i=2}^{M-1} \sum_{j=2}^{N-1} [S(i,j) - \bar{S}]^2} \quad (6)$$

式中: $S(i,j)$ 为地磁坡度值; $\bar{S}$ 为该区域地磁坡度均值。

#### 7) 地磁相关系数 $\rho$ 。

地磁相关系数表征了地磁数据的独立性,其值越小,越有利。可表示为<sup>[9]</sup>

$$\rho = (\rho_x + \rho_y)/2 \quad (7)$$

式中: $\rho_x = \frac{1}{M(N-1)\sigma^2} \sum_{i=1}^M \sum_{j=1}^{N-1} [f(i,j) - \bar{f}][f(i,j+1) - \bar{f}]$ ; $\rho_y = \frac{1}{(M-1)N\sigma^2} \sum_{i=1}^{M-1} \sum_{j=1}^N [f(i,j) - \bar{f}][f(i+1,j) - \bar{f}]$ 。

## 1.2 基于PCA的地磁特征参量的选择

上述地磁特征在不同程度上反映了地磁场的某些特征,由于参量间可能存在共线性问题,神经网络又是以地磁图特征作为输入,进而会潜在地导致神经网络计算结果存在较高的预报方差和病态<sup>[11]</sup>,因此地磁图特征的选择问题也应该被考虑。

主成分分析法(PCA)主要是采用降维的思想进行特征选择,即将原来多数属性通过线性变换为少数几个保留原始信息且不相关的综合属性。具体步骤如下。

1) 属性同趋势化:鉴于上述特征中既有高优指标,又有低优指标,首先应对各属性同趋势化处理。算式为 $x'_{ij} = -x_{ij}$ ,其中, $x_{ij}$ 为低优指标。进而得到的新特征矩阵记为 $\mathbf{X}' = [x'_{ij}]_{m \times n}$ ( $m$ 为候选匹配区个数, $n$ 为特征参量个数)。

2) 标准化处理:由于上述地磁特征值并不是同一个数量级别,且各特征单位不统一,这样容易出现奇异数据,因此标准化处理是必须的,且不会减弱原有数据的差异。算式为

$$z_{ij} = \frac{x'_{ij} - \bar{x}_j}{\sigma_j} \quad (8)$$

式中: $\bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij}$ ; $\sigma_j = \frac{1}{m-1} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2$ 。标准化处理后的矩阵记为 $\mathbf{Z} = [z_{ij}]_{m \times n}$ 。

3) 对标准化矩阵  $Z$  求相关矩阵  $R$ , 为

$$R = [r_{ij}]_{n \times n} \quad r_{ij} = \frac{1}{m} \sum_{i=1}^m z_{ik} z_{j} \quad (9)$$

4) 解得矩阵  $R$  的特征根系, 将得到的特征根  $\lambda_g$  按照从大到小的顺序排列, 其大小反映了各主成分的作用大小, 并求得特征向量  $L_g$ 。方差贡献率算式为

$$p_g = \lambda_g / \sum_{g=1}^n \lambda_g \quad (10)$$

通过累积贡献率来确定  $k$  值, 为了确保准确率, 设判定准则为  $\sum_{g=1}^k p_g \geq 99\%$ 。

5) 计算结果矩阵  $T = ZL$ , 其中,  $L$  为特征矩阵;  $T_1$  为第 1 主成分,  $T_2$  为第 2 主成分,  $\dots$ ,  $T_n$  为第  $n$  主成分, 并只取其前  $k$  个主成分。

在某区域背景场内, 选取 250 块格网大小为  $15 \times 15$ 、精度为 200 m/格的子区, 提取 1.1 节中提到的 7 个地磁特征参量, 按照上述步骤进行主成分分析, 得到每个主成分对应的特征值和贡献率如表 1 所示。

表 1 每个主成分对应的特征值和贡献率

Table 1 Eigenvalues and contribution rates of each principal component

	序号 $g$						
	1	2	3	4	5	6	7
特征值 $\lambda_g$	5.2457	1.0133	0.7023	0.2236	0.0471	0.0124	0.0052
贡献率 $p_g$	0.7236	0.1398	0.0969	0.0308	0.0065	0.0017	0.0007
累积贡献率	0.7236	0.8634	0.9603	0.9911	0.9975	0.9993	1.0000

从表 1 可直观看出, 原数据 99.75% 的信息量均集中在前 5 维主成分, 因此在神经网络中主要输入前 5 个主成分作为独立参量。

## 2 地磁适配区选择方法设计

### 2.1 基于 GA-BP 神经网络的地磁适配/非适配区划分

鉴于传统 BP 神经网络修正权值和阈值是采取梯度下降的方式, 有容易陷入局部极小值的缺陷, 因此本文对 BP 神经网络用遗传算法(GA)进行优化。

GA-BP 神经网络模型是通过在 BP 神经网络训练过程中引入遗传算法, 以网络权值和阈值作为种群个体来仿照基因编码, 使用样本观测值和预测值的误差绝对值之和作为个体适应度函数, 通过选择、交叉、变异操作的不断迭代进化, 最终得到种群最优个体, 并通过解码, 得到 BP 神经网络全局最优权值和阈值, 从而建立起的模型。

在 1 节地磁特征参量的提取与选择的基础上, 通过对 GA-BP 神经网络的训练, 形成特征参量和适配性标签之间的网络映射。以候选适配区的地磁特征参量作为神经网络的输入量, 利用神经网络进行分类并获

得相应的适配性标签, 以确定适配区, 主要通过两个阶段来构建: 1) ~ 3) 为网络构建和训练阶段; 4) 为测试阶段, 即网络应用。具体步骤如下。

1) 在某区域地磁图上, 以一定的候选适配区作为研究对象, 提取出 7 个地磁特征参量, 然后用 PCA 对其进行分析, 最终得到 5 个新的独立统计参量。

2) 借助相关匹配算法, 本文选用均方差算法(MSD), 在候选地磁适配区域进行匹配定位试验, 以匹配概率作为适配性指标, 根据精度要求, 当匹配概率大于等于 0.9 时, 记为 1, 表示该样本为适配区; 当小于 0.9 时, 记为 0, 表示该样本为非适配区。1 和 0 即为适配标签。

3) 将 1) 中的独立统计参量作为训练样本的输入, 2) 中对应的适配性标签作为训练样本的输出。采用 GA 优化 BP 神经网络参数, 再对其进行训练, 进而获得稳定的分类模型。

4) 计算给定区域的地磁特征参数, 并对其进行 PCA 分析, 获得独立的特征参量, 随后输入到 3) 中训练好的分类模型中, 即可实现对该区域适配/非适配性能的识别和分类。

### 2.2 分类能力评价指标

利用上述方法进行分类时, 可能会产生两种错误: 1) 将适配区( $C1$ ) 误断为非适配区( $C2$ ); 2) 将非适配区( $C2$ ) 误断为适配区( $C1$ )。

错误 1) 的概率可以表示为

$$P_1 = \frac{N(\text{分类结果} \in C2 \cap \text{实际类别} \in C1)}{N(\text{实际类别} \in C1)} \quad (11)$$

错误 2) 的概率为

$$P_2 = \frac{N(\text{分类结果} \in C1 \cap \text{实际类别} \in C2)}{N(\text{实际类别} \in C2)} \quad (12)$$

另外, 本文用分类精度作为分类性能的评价指标。其算式为

$$\text{分类精度} = \frac{N(\text{分类结果} \in C1 \cap \text{实际类别} \in C1) + N(\text{分类结果} \in C2 \cap \text{实际类别} \in C2)}{N(\text{实际类别} \in C1) + N(\text{实际类别} \in C2)} \quad (13)$$

很明显地, 当犯两种错误类型(特别是错误 2)) 的概率较小, 且分类精度较高时, 分类性能较为理想。

### 2.3 神经网络参数的确定

使用 GA-BP 神经网络进行地磁适配区/非适配区的识别分类时, 设置遗传算法的最大进化代数为 100, 种群大小为 50。染色体选择策略为轮盘赌; 其交叉策略为单点交叉, 交叉概率为 0.6; 变异策略为随机变异, 变异概率为 0.05; 适应度函数为 BP 神经网络的误差。

其次设置 BP 神经网络输入层为 5 个神经元, 而输出分为 2 类, 可以用 1 位二进制数来表示。因此, 输出层含有 1 个神经元。此外, 通过对神经网络的训练来

确定隐层节点的个数,结果如图1所示。

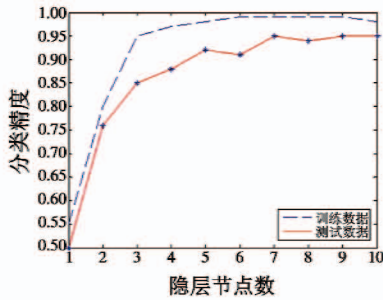


图1 分类精度随隐层节点数变化的曲线

Fig.1 The change of the classification accuracy with the number of hidden layer nodes

从图1可看出,当使用7个神经元时,分类精度较高,因此,隐层节点个数选为7。

### 3 仿真验证与分析

为了验证上述方法的正确性,用1节中在某区域背景场内选取的250块网格大小为15×15、精度为200 m/格的子区作为实验对象。然后对各候选适配区开展相关匹配实验,计算匹配概率,对候选适配区进行分类划分,给出适配/非适配标签,并计算各子区7个地磁特征参数及其主成分,在250块候选适配区中随机选择125块区域作为训练样本,其余125块区域作为测试样本。

#### 3.1 仿真验证

本文将主成分分析法与GA-BP神经网络结合,从特征选择和参数优化两个角度来提高候选适配区的分类精度,为验证该方法的优越性,通过以下4组实验来分析:1)不进行特征选择,直接利用BP神经网络;2)不进行特征选择,但利用遗传算法优化;3)利用PCA进行特征选择,无参数优化;4)利用PCA进行特征选择,并且进行参数优化。实验结果如表2所示。

表2 4组分类方法的性能比较

Table 2 Performance of four groups of classification methods

分类方法	分类性能值		分类精度
	$P_1$	$P_2$	
无+BP	0.232 5	0.035 7	0.903 2
无+GA-BP	0.223 6	0.031 7	0.922 5
PCA+BP	0.237 4	0.028 5	0.926 1
PCA+GA-BP	0.196 5	0.021 2	0.960 4

由表2可以看出,本文提出的PCA和GA-BP神经网络相结合的方法,在完成特征选择和参数优化的基础上,能够有效地提高候选适配区的分类精度。

#### 3.2 匹配验证

随意选取3块大小为15×15、网格精度为200 m/格的候选适配区M1,M2和M3,分别计算其匹配概率,并利用本文的方法对其进行分类。结果如表3所示。

表3 候选适配区M1~M3的分类结果

Table 3 Classification results of candidate matching area M1 to M3

候选适配区	匹配概率	分类结果
M1	0.942 5	适配区
M2	0.795 2	非适配区
M3	0.657 0	非适配区

由表3可知,GA-BP神经网络的分类结果,与该区域实际的适配性能(匹配概率)有很好的 consistency,即分类结果为适配区的区域(M1)对应的匹配概率也高,分类结果为非适配区的区域(M2,M3)匹配概率相对较低,可以看出该结果可信,因此可以将M1选作适配区。

图2为这3块区域的相关匹配图,表4为区域M1~M3的平均定位精度。从表4看出,在利用本文所述方法选出的适配区(即M1)内进行匹配定位,精度比较理想。

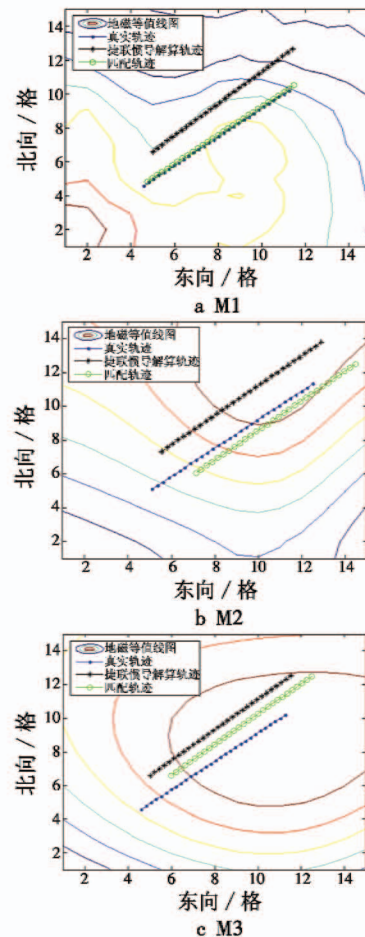


图2 区域M1~M3的相关匹配图

Fig.2 Correlation matching graph of area M1 to M3

表4 区域M1~M3的平均定位精度

Table 4 Average positioning accuracy of area M1 to M3

	候选适配区		
	M1	M2	M3
平均定位精度	(70,120)	(380,220)	(530,550)

#### 4 结束语

为了使候选匹配区的分类精度得到提升,本文提出了基于 PCA 和 GA-BP 神经网络结合的地磁匹配区的选取方法。首先,利用主成分分析法进行候选匹配区的特征选择,该方法在一定程度上解决了由于参量间的共线性导致的神经网络计算结果存在预报方差和病态的问题,同时也减少了计算量。其次,进一步对 BP 神经网络的参数通过 GA 进行优化,更加提升了分类的准确性。经验证,本文方法分类精度较高,选择出来的适配区域能有效地提高地磁导航定位精度。在实际应用中,能够正确地选取匹配性最优的区域作为匹配区,对下一步进行匹配定位有重要的意义。

#### 参考文献

- [1] 李婷,张金生,王仕成,等. 量子粒子群算法在地磁匹配航迹规划中的应用[J]. 电光与控制,2015,22(7):43-47.
- [2] 陈丰刚. 地磁辅助导航适配区选取方法研究[D]. 长沙:国防科学技术大学,2013.
- [3] YUAN S, ZHANG J S, QIAO Y K, et al. Matching area intelligent selection method in geomagnetic navigation [C]//International Conference on Intelligent Computing and Intelligent Systems, IEEE, 2010:860-864.
- [4] 罗海波,常铮,余新荣,等. 采用多特征融合的自动适配区选择方法[J]. 红外与激光工程,2011,40(10):2037-2041.
- [5] WANG P, WU Y, HU X, et al. Geomagnetic aided navigation suitability evaluation based on principal component analysis [C]//International Conference on Industrial Control and Electronics Engineering, 2012:324-329.
- [6] WANG K, LI Y, RIZOS C. Research on terrain suitability of iterative closest contour point algorithm for underwater navigation [C]//The 22nd International Technical Meeting of the Satellite Division of the Institute of Navigation, 2009:866-870.
- [7] 吕云霄,吴美平,胡小平. 基于支持向量机的地磁辅助导航匹配区域选取准则[J]. 兵工自动化,2011,30(1):49-52.
- [8] 李婷,张金生,王仕成,等. 基于改进地磁熵的地磁适配区选择准则[J]. 大地测量与地球动力学,2014,34(5):151-155.
- [9] 胡小平. 水下地磁导航技术[M]. 北京:国防工业出版社,2013.
- [10] 康崇,王明,樊黎明,等. 基于地磁熵和地磁差异熵的地磁导航区域选取[J]. 应用基础与工程科学学报,2015(6):1156-1165.
- [11] 张凯,赵建虎,施闯,等. BP 神经网络用于水下地形适配区划分的方法研究[J]. 武汉大学学报:信息科学版,2013,38(1):56-59.
- (上接第 97 页)
- [5] 袁磊,孙永荣,周晓达,等. 基于 ARINC661 的 DF 文件设计平台研究与实现[J]. 现代电子技术,2012,35(5):156-159.
- [6] 邢飞燕,施佳,胡祖辉,等. 基于 MVC 的高校研究生会管理系统的开发与实现[J]. 科技资讯,2015,13(18):34-35.
- [7] 牛文生. 机载计算机技术[M]. 北京:航空工业出版社,2013:19-25.
- [8] 赵文献,孙永荣,王勇,等. ARINC661 设计文件自动生成平台的研究与实现[J]. 航空计算技术,2014(2):114-117.
- [9] 赵文献. ARINC661 定义文件开发软件的设计与实现[D]. 南京:南京航空航天大学,2016.
- [10] 郑澄,孙永荣,王晖,等. 基于 ARINC661 的触摸屏显控系统研究与设计[J]. 电光与控制,2015,22(4):62-65.
- (上接第 109 页)
- 雷达识别技术[J]. 中国电子科学研究院学报,2009,4(6):589-593.
- [3] 罗长胜,吴华,程嗣怡. 脉冲幅度特性分析的相控阵雷达识别技术[J]. 火力与指挥控制,2013,38(7):27-31.
- [4] 程胜,王星,楚维,等. 基于脉冲幅度特性的相控阵雷达识别技术研究[J]. 计算机工程与应用,2012,48(23):120-124,211.
- [5] 叶巍,牟连云,李仙茂. 基于脉冲包络的相控阵雷达识别技术研究[J]. 航天电子对抗,2011,27(1):41-44.
- [6] 冀琛,潘谊春,郁春来,等. 基于脉幅信息的相控阵体制雷达快速识别[J]. 空军预警学院学报,2013,27(5):323-326.
- [7] 李欣,王春阳,王强,等. 基于 Hausdorff 距离的相控阵雷达波束指向侦察方法研究[J]. 中国电子科学研究院学报,2013,8(5):507-511.
- [8] 尹亮,姜秋喜,潘继飞,等. 一种相控阵雷达指纹特征提取技术[J]. 航天电子对抗,2012,28(5):47-49.