

# 基于遗传算法的 SAR 图像目标鉴别特征选择

高 贵<sup>1</sup>, 周蝶飞<sup>2</sup>, 张 军<sup>1</sup>, 黄纪军<sup>3</sup>, 李德仁<sup>4</sup>

(1. 国防科技大学信息系统与管理学院, 湖南长沙 410073; 2. 湖南师范大学, 湖南长沙 410081;

3. 国防科技大学电子科学与工程学院, 湖南长沙 410073; 4. 武汉大学, 湖北武汉 430089)

**摘 要:** 针对 SAR 图像自动目标鉴别的应用, 提出了一种基于遗传算法的特征选择方法. 首先提取了反映目标和杂波虚警差异的八个特征, 分别是: 四个空间边界属性特征, 一个分形维数特征和三个对比度特征. 然后对由八个特征构成的特征矢量采用遗传算法进行特征选择, 以选出对于目标鉴别最优的特征序列. 遗传算法中适应度函数的设计综合考虑了描述长度、鉴别总错误数以及漏报数等三个因素, 使得该适应度函数对于特征优劣的评价更全面. 实测数据的实验结果证明了所提算法的有效性.

**关键词:** 合成孔径雷达; 鉴别; 特征提取; 特征选择

**中图分类号:** TN957 **文献标识码:** A **文章编号:** 0372-2112 (2008) 06-1042-06

## Feature Selection for Target Discrimination in SAR Images Based on Genetic Algorithm

GAO Gui<sup>1</sup>, ZHOU Diefei<sup>2</sup>, ZHANG Jun<sup>1</sup>, HUANG Jijun<sup>3</sup>, LI De2ren<sup>4</sup>

(1. School of Information System and Management, NUDT, Changsha, Hunan 410073, China;

2. Hunan Normal University, Changsha, Hunan 410081, China;

3. School of Electronic Science and Engineering, NUDT, Changsha, Hunan 410073, China;

4. Wuhan University, Wuhan, Hubei 430089, China)

**Abstract:** For the application of target discrimination in SAR images, a method of feature selection based on Genetic Algorithm (GA) is proposed in this study. First, eight features including four spatial edge property features, one fractal dimension feature and three contrast features, are extracted to distinguish target and clutter false alarm. Then, the GA is applied to select the best feature series from the feature vector consisting of eight features. The fitting function in GA is devised by considering three factors including describing length, the total wrong number of discrimination and losing number. So it can evaluate the feature better than other algorithms. The experimental results of actual data demonstrate the presented algorithm is effective.

**Key words:** synthetic aperture radar; discrimination; feature extraction; feature selection

## 1 引言

SAR 图像目标鉴别是 SAR ATR 研究领域的一个子课题<sup>[1]</sup>. 其主要目的是对目标检测得到的潜在目标区域进行进一步处理, 以去除自然杂波虚警和部分人造杂波虚警, 从而减小目标分类/识别的代价. SAR 图像自动目标鉴别实现的一般步骤是<sup>[2-3]</sup>: 首先提取一些有利于目标鉴别的特征, 构成鉴别的特征序列; 然后根据不同的鉴别应用, 对提取的特征序列进行特征选择, 以选出最优的特征组合; 最后采用鉴别器完成目标的鉴别决策.

特征选择是影响目标鉴别性能的关键步骤<sup>[1]</sup>. 实际中, 由于特征的相关性造成的冗余, 噪声的存在等都会

影响鉴别性能, 而且不同雷达成像条件下鉴别采用的最优特征序列有所不同, 因此特征选择在目标鉴别实现中是必要的. 目前的特征选择大多来源于专家的经验, 然而, 随着问题的深入以及数据源的丰富, 即使是最有经验的专家也无法指出, 具体到某一鉴别应用, 什么样的特征组合最优<sup>[4-6]</sup>.

解决特征选择问题的核心在于<sup>[7]</sup>: 必须设计一种对于不同鉴别应用都适用, 并具有较强的从中搜索得到最佳特征子序列能力的特征选择方法. 遗传算法<sup>[5-7]</sup> (GA) 由于本身是一种速度较快、性能较好的全局优化方法, 在近几年的鉴别特征选择中取得了成功. 但是目前提出的基于遗传算法的鉴别特征选择方法的主要

问题<sup>[5-7]</sup>是:在适应度函数的设计中,仅仅考虑了两个因素:被选特征个数和总错误率,来评估被选特征子序列的优劣.这对所选特征优劣的评价是不全面的,容易造成得到的最优特征子序列应用于鉴别时,目标的漏报增多或目标鉴别的精度下降.

本文借鉴 Bhanu 小组<sup>[5,6]</sup>提出的基于遗传算法的特征选择方法并进行改进,提出一种对特征优劣评价更全面的特征选择方法.本文基于 GA 的特征选择的实现思路是:以从训练图像集提取的特征序列集为输入,通过改进遗传算法的适应度函数,经过 GA 的最优搜索,得到最优的特征子序列.第二节提取了目标鉴别的特征;第三节给出了基于 GA 的特征选择方法;第四节为实验结果与分析;最后为结论部分.

## 2 目标鉴别特征提取

### 2.1 空间边界属性特征提取<sup>[7]</sup>(特征 1~4)

空间边界属性特征被认为是一类对于目标鉴别非常有用的特征.理论上,空间边界属性特征可以提供目标或杂波的尺寸、形状、结构等诸多信息.目标或杂波虚警的空间边界属性可以通过其对应图像的边界属性近似加以表征.

设一幅待鉴别的强度图像为  $I$ , 大小为  $N \times M$ , 在位置  $(I, J)$ , 像素值是  $I_{ij}$ , 那么图像可以表示为  $I = \{I_{ij}, 1 \leq i \leq N, 1 \leq j \leq M\}$ . 如果图像相对于某一个参数变化, 则图像记为  $I(t)$ , 相应的像素记为  $I_{ij}(t)$ . 预处理的过程是:首先,对输入的强度图像取对数运算,以增加图像的动态范围.然后计算图像的中值  $I$ , 每个像素值减去中值得  $I$  到规范化图像,记变换后的图像像素

$$x_{ij} = I_{ij} - I.$$

为了展现目标区域高强度像素的聚类,从而最好地获取目标/信号0,寻找图像的最优阈值是必要的.对经过预处理的图像切片采用下式的 sigmoidal 函数进行变换:

$$s(x; t) = \frac{1}{1 + \exp(-c(x - t))}, -\infty < x < +\infty \quad (1)$$

其中  $x$  为强度图像,  $t$  为图像的阈值,  $c$  为一个正实数,表征 sigmoidal 函数的陡峭度.经过 sigmoidal 函数变换后的图像像素值在 0~1 之间.

原图像  $x$  通过 sigmoidal 函数变换后,能够产生一族随着阈值参数  $t$  变化的新的图像  $w(t)$ , 其中的每一个像素值为:

$$w_{ij}(t) = s(x_{ij}; t) \quad (2)$$

通过观察在整个阈值  $t$  的变化范围内  $x(t)$  的变化,能够了解图像切片内含有的目标或杂波的结构,获得大量的有用信息.为了提取这些信息,必须度量在某个阈值被选取的情况下,图像是否表现出了一个一致的高强

度的像素的聚类.一种简单的度量方法是图像中所含目标或杂波虚警像素空间分布的离散程度.图像像素质量关于图像质心位置空间离散程度的度量可由下式表示:

$$r(t) = \frac{\sum_{i=1}^N \sum_{j=1}^M w_{ij}(t) \sqrt{(i - i_c(t))^2 + (j - j_c(t))^2}}{\sum_{i=1}^N \sum_{j=1}^M w_{ij}(t)} \quad (3)$$

其中  $i_c$  和  $j_c$  是图像的质心坐标.当  $t$  的值很低时,那么图像中几乎所有的像素都被认为是活动的,其像素强度值接近为 1,  $r(t)$  的取值接近它的取值上限.此时,对于阈值  $t$  中等程度的增加或者减少,的值几乎不变.

如前所述,很自然的把图像中所能提供的目标信息量和统计量  $r(t)$  的幅度变化联系在一起,即统计量的一阶导数  $|rc(t)| \cdot |rc(t)|$  的最大值表示了统计量变化的最大值,即所能给出的最大信息量所对应的阈值,定义  $t_0$  为对应的阈值,则:

$$t_0 = \underset{t \in T}{\operatorname{argmax}} \{ |rc(t)| ; r^n(t) = 0 \} \quad (4)$$

$T$  为实际中阈值参数的变化范围.如果在阈值参数区间内,没有二阶导数为零的阈值点,则表达式应为:

$$t_0 = \underset{t \in T}{\operatorname{argmax}} \{ |rc(t)| \} \quad (5)$$

一旦最优的阈值  $t_0$  被确定,就可以提取相应特征.

确定最优阈值  $t_0$  后,提取的空间边界属性特征如下:

# 图像像素质量的平均值:

$$f_1 = \log \frac{\sum_{i=1}^N \sum_{j=1}^M w_{ij}(t_0)}{MN}$$

# 图像像素空间离散程度特征:  $f_2 = r(t_0)$ , 表征了图像亮像素相对于其质心的加权平均距离,权值正比于像素的质量;

# 拐点特征:  $f_3 = rc(t_0)$ , 它表征了在最优阈值上,图像中像素空间离散程度的改变量;

# 加速度特征:  $f_4 = r\hat{c}(t_0)$ , 它测量在最优阈值下像素质量空间离散程度改变率的加速度.

### 2.1.2 分形维数特征提取<sup>[8]</sup>(特征 5)

分形维数是一个表征目标或杂波虚警空间维数的物理量,它度量了二值图中强点排列的紧密程度.因为感兴趣目标的强点的排列一般较自然杂波虚警像素的排列要更紧密,所以可以用来区别目标和杂波虚警.该特征定义为在潜在目标切片图像中最亮点空间分布的 Hausdorff 维数:

$$f_5 = \frac{\log N_1 - \log N_2}{\log d_2 - \log d_1} \quad (6)$$

其中  $N_1$  是能够覆盖强点的  $1 \times 1$  pixel 大小的最小窗口数,  $N_2$  是能够覆盖强点的  $2 \times 2$  pixel 大小的最小窗口数,  $d_1 = 1$ ,  $d_2 = 2$  为窗口大小.

### 21.3 对比度特征的提取(特征 6~8)

对于感兴趣目标而言, 主要由金属制成, 其后向散射比较强, 在图像上表现为相对于背景较强的对比度. 因此, 对比度特征从理论上也可以作为目标鉴别的一类特征.

在确定了潜在目标切片图像的最优阈值  $t_0$  后, 把大于最优阈值的像素认为是/ 活动像素0, 其余像素为/ 背景像素0, 那么对比度可以由下面 3 个特征加以刻画:

均值信噪比特征:

$$f_6: M_{SNR} = \frac{1}{N_T} \sum_{i=1}^{N_T} \left( \frac{m_i}{R} \right) \quad (7)$$

其中  $N_T$  为/ 活动像素0 的个数,  $m_i$  为/ 活动像素0 强度值,  $R$  为/ 背景像素0 均值. 平均信噪比反映了图像切片内所含目标或杂波虚警相对于背景的平均对比度.

峰值信噪比特征:

$$f_7: P_{SNR} = \frac{\max(m_i, i \in \{1, 2, \dots, N_T\})}{R} \quad (8)$$

峰值信噪比反映了图像切片内所含目标或杂波虚警相对于背景的最大对比度.

强点百分比特征:

$$f_8: P_{\epsilon_{SNR}} = \frac{1}{N_T} |\{m_i | m_i > P_{SNR} \# p\%, 1 \leq i \leq N_T\}| \quad (9)$$

其中  $|\#|$  表示集合的势, 该式即计算对比度超出峰值信噪比  $p\%$  的强点像素占有/ 活动像素0 的比例. 在后续的实验中, 选择  $p = 50$ , 即把峰值信噪比下降 3dB 的值作为阈值.

## 3 基于遗传算法的特征选择

### 3.1 基于 GA 的特征选择

遗传算法是模拟生物在自然环境中的遗传和进化过程而形成的一种自适应全局优化概率搜索算法. 遗传算法是操作在二进制串搜索空间上的一个最优化的过程. 搜索空间中的每个点是由 0 和 1 组成的一个有限序列表示的, 其中 0 或 1 叫做一个基因, 对应的由 0-1 码组成的二进制串称为染色体或个体, 由一定数量个体形成的集合称为种群或群体.

尽管遗传算法的不同实现在细节上有所不同, 但一般需要解决下列问题: 染色体编码、初始种群的设定、适应度函数的设计、遗传算子的确定、终止条件设定等. 具体到本节的特征选择问题, 各步如下:

(1) 定义染色体

特征选择的输入是从训练图像提取的初始特征序列. 若初始特征序列中包含几个特征, 则定义染色体为长度是 0-1 的二进制码, 如图 1 所示, 对应于前面提取的特征, 本文的特征总数. 图中特征序号分别对应于上

文的特征 1~8. 如果染色体的第位为 1 则表示对应的特征被选中, 为 0 则表示对应的特征未被选中. 每一个染色体对应于初始特征序列的一个特征子序列.

特征序号	1	2	3	4	5	6	·	·	·	l
染色体	1	0	1	0	0	0	·	·	·	1

图 1 特征序列染色体编码

(2) 初始种群的设定

随机生成长度为 1 的染色体  $N$  个作为初始种群.  $N$  为种群的大小. 为了保证种群中个体的多样性, 本文中取  $N = 100$ .

(3) 设计适应度函数

适应度函数的设计被用来评估当前种群中每一个特征子序列的性能. 特征子序列性能的评估需要综合考虑以下因素:

**描述长度:** 特征子序列可以认为是描述数据的模型, 因此描述长度定义为特征子序列中特征的个数, 即对应染色体中编码为 1 的基因个数.

**总错误数  $n_e$ :** 目标鉴别是一个两类模式识别问题. 采用二次距离鉴别器<sup>[2,3]</sup> 计算鉴别的总错误数. 假设特征子序列中特征的联合分布满足高斯分布, 则二次距离鉴别器理论上能给出最优的鉴别性能. 设目标类为  $C_1$ , 杂波类为  $C_2$ , 由  $C_i$  中训练图像提取的对应于特征子序列的特征矢量为  $f_{i1}, f_{i2}, \dots, f_{iM}$ ,  $M$  为  $C_i$  中训练图像的个数. 那么类  $C_i$  的均值和协方差矩阵由下式计算:

$$L_i = \frac{1}{M} \sum_{j=1}^M f_{ij}, \quad 2 = \frac{1}{M} \sum_{j=1}^M (f_{ij} - L_i)(f_{ij} - L_i)^T \quad (10)$$

对于每一幅训练图像, 若提取的对应于特征子序列的特征矢量为  $f$ , 则该特征矢量到类  $C_i$  的二次距离为:

$$d_i = (f - L_i)^T 2_i^{-1} (f - L_i) \quad (11)$$

对于训练图像库中的每一幅图像, 若计算的  $d_1 < d_2$ , 则认为该图像是目标图像, 反之, 则认为该图像是杂波虚警. 因此总错误数  $n_e$  定义为目标误判为杂波虚警的个数和杂波虚警误判为目标的个数之和.

**漏报数  $m_e$ :** 设训练数据库中目标类的数据个数为  $m$  个, 则漏报数可定义为:

$$m_e = |\{ic | d_{e1} > d_{ic2}, 1 \leq ic \leq m\}| \quad (12)$$

其中  $|\#|$  表示集合的势,  $d_{e1}$  是第  $ic$  个目标图像到类  $C_1$  的二次距离,  $d_{ic2}$  是第  $ic$  个目标图像到类  $C_2$  的二次距离. 上式即表示训练数据库中目标图像误判为杂波虚警图像的个数.

适应度函数是综合考虑上述三个因素形成的:

一方面, 由于待评估的特征子序列可看作是描述训练数据的模型, 那么特征子序列中所含特征数越多 (即描述长度越大) 则意味着近似训练数据的模型越复杂. 势必引起大的计算代价, 并有可能过于敏感于数据的特

性,使得随机噪声也被模型化,引起过拟合问题;如果特征子序列的描述长度小,那么它可能不能充分描述实际数据的特性,对应的鉴别性能不能达到要求。

另一方面,特征子序列的鉴别精度也是衡量其优劣的重要指标。在训练数据库中数据总个数和目标数据个数已知的情况下,鉴别精度可以由总错误数和漏报数确定。总错误数对应于总错误率,反映了用特征子序列进行目标鉴别的总体精度。总错误数越少,鉴别的总体精度越高。然而,总错误数把目标误判为杂波虚警以及杂波虚警误判为目标两种错误情况置于同等地位,并不能完全表征特征子序列的鉴别精度。Bhanu 研究小组<sup>[5,6]</sup>在特征子序列的鉴别精度方面只考虑了总错误数,这是不全面的。因为:事实上,我们更关心的是目标图像尽量不发生漏报,而允许杂波虚警存在误判。换句话说,我们希望在一定的总体鉴别精度下,漏报数越少越好。

上述的三个因素是相互制约的。因此,评价特征子序列的优劣,必须在上述三个因素中折衷。

从特征子序列的实际应用考虑,自然地,可以认为鉴别性能好且具有最小描述长度的特征子序列是最优的。因此,定义适应度函数为:

$$F(f) = - (k \log_{10}(1) + n_e \log_{10}(n) + m_e \log_{10}(m)) \quad (13)$$

其中  $f$  是提取的对应于特征子序列的特征矢量,  $k$  是描述长度,对应特征子序列染色体中编码为 1 的基因的个数。1 表示特征总数,即染色体长度。 $n_e$  是总错误数,  $n$  是训练图像总数。 $m_e$  是漏报数,  $m$  是训练数据库中目标数据总数。

很容易看出特征子序列中被选择的特征数越少、误分类的错误数越小且误分类数据中漏报数越少,适应度函数的值越大。

#### (4) 确定遗传算子

在 GA 中,个体的进化是在遗传算子的作用下完成的。最主要的遗传算子有选择算子、交叉算子和变异算子。本文中三种算子的选取和文献<sup>[5,6]</sup>相同。初始交叉率  $Q=0.18$ ,初始变异率  $Q_m=0.101$ 。为了减少训练时间同时防止种群早熟,在遗传算法的搜索过程中,如果最好个体的适应度值超过 3 代仍然没有提高,那么 GA 以 0.102 的幅度按代增加变异率。一旦最好个体的适应度增加,则重置变异率  $Q_m=0.101$ 。

#### (5) 设定终止条件

本算法的终止条件设定为:当遗传算法的迭代次数等于最大迭代次数 MaxGen,或变异率大于 0.109。本算法设定 MaxGen=50。

### 3.1.2 基于 GA 的特征选择算法流程

算法流程具体为(其中为迭代次数):

# 初始化种群:  $P_z$  随机产生  $N$  个长度为 1 的染色体(个体);

```
# while( Gen [ MaxGen 或  $Q_m$  [ 0.109 ] ), do
11 适应度评价: 对  $P$  中每个个体  $h$ , 计算  $F(h)$ ;
21 产生新一代种群  $P_s$ ;
if 最优个体适应度超过 3 代仍没有增加)
    1) 在原有变异率基础上以 0.102 幅度增加变异率;
    2) 选择、交叉、变异;
    3) 种群更新:  $P_z \leftarrow P_s$ ;
else
    1) 置变异率:  $Q_m=0.101$ ;
    2) 选择、交叉、变异;
    3) 种群更新:  $P_z \leftarrow P_s$ ;
end
end
# 从  $P$  中返回最优个体
```

## 4 实验结果与分析

### 4.1 实测数据结果

采用 MSTAR 数据库进行实验。MSTAR 数据的名义分辨率是 0.13m@0.13m,极化方式:HH 极化,工作波段:X 波段,成像方式:聚束模式。其中目标图像库采用 17°俯仰角下, BIR70 装甲车、T72 坦克、BMP2 装甲车的图像共 233 幅,这些图像分布在整个 360°方位角范围内,图像大小为 128@128。为使提取的杂波虚警图像更相似于目标,设定低虚警概率  $p_{fa}=0.1002$ 。对纽约 redstone 地区的 16 幅杂波图像(共计覆盖区域大小为 116323km<sup>2</sup>)采用文献<sup>[7]</sup>提出的检测器进行检测并聚类,得到 538 幅 128@128 大小的杂波虚警图像作为杂波虚警图像库。

按照既定的方案,实验的过程可简述为:首先,随机选择 150 幅目标图像和 246 幅杂波虚警图像作为训练数据库,其余的 83 幅目标图像和 292 幅杂波虚警图像作为测试数据库。然后,对每一幅训练图像提取前述的 8 个特征,得到对应的特征序列,进而把由整个训练数据库得到的特征序列集采用基于 GA 的特征选择方法进行特征选择,得到最优的特征子序列。

采用训练数据库经本文遗传算法进行特征选择。若以初始种群为第 0 代,则独立进行 10 次基于 GA 的特征选择,结果选出的最优特征子序列都是  $[f_1, f_2, f_3, f_4, f_8]$ ,对应的适应度值都是 -711131。10 次独立特征选择实验算法终止时的平均迭代次数(或代数)为 91.7,平均耗时 301.7698s,这说明了本文遗传算法进行特征选择的速度很快。

表 1 给出了本文算法选出的最优特征序列下采用式(10)的目标鉴别结果。表 2 给出了利用所有的 8 个特征进行目标鉴别时的鉴别精度。比较表 1 和表 2,可以看出,所有特征对于训练数据的鉴别精度低于最优特征

子序列的鉴别精度. 因此提取的最优特征子序列应用于目标鉴别是有效的.

对于测试数据库, 采用所有特征获得的鉴别精度则较训练数据的鉴别精度降低较大. 采用所有特征对测试数据库进行鉴别所获得的精度要低于对训练数据库所获得的鉴别精度, 并且采用所有特征对测试数据的鉴别精度明显低于最优特征子序列下对测试数据目标鉴别的精度.

41.2 与已有算法的对比

基于 GA 的目标鉴别特征选择方法首先是由 Bhanu 研究小组<sup>[5 6]</sup>提出的. 该方法与本文特征选择算法的不同之处是适应度函数的设计上. Bhanu 研究小组特征选择方法的适应度函数定义为<sup>[5 6]</sup>:

$$F_B(f) = - (k \log_{10}(l) + n_e \log_{10}(n)) \quad (14)$$

其中  $k$  是描述长度, 对应特征子序列染色体中编码为 1 的基因的个数.  $l$  表示特征总数, 即染色体长度.  $n_e$  是总错误数,  $n$  是训练图像总数. 即该适应度函数在特征子序列的鉴别精度方面只考虑了总错误数. 在遗传算法的迭代过程中, 考虑下列两个特征子序列(种群中的两个个体)编码:  $[0\ 1\ 1\ 1\ 1\ 0\ 0\ 1]$  和  $[1\ 1\ 1\ 1\ 1\ 0\ 0\ 0]$ , 其用于目标鉴别的精度如表 3 所示.

由表 3 可知, 由于两个特征子序列的被选特征个数相同, 而且对于训练数据的总错误数也相同. 因此在 Bhanu 算法中, 两者的适应度值相同, 都为 - 917108, 也就是说: Bhanu 算法认为这两组特征用于目标鉴别时其性能是相当的. 但是明显地, 第一个特征子序列对于训练数据产生的鉴别错误全是漏报, 而第二个特征子序列对于训练数据产生的鉴别错误中只有一个漏报. 由于我们在目标鉴别时更关心的目标不被漏报, 因此从理论上讲, 第二个特征子序列要优于第一个特征子序列. 这一点在 Bhanu 算法中却无法体现, 恰恰说明了仅仅考虑了被选特征个数和总体错误率来评估被选特征子序列的优劣是不全面的.

应用本文特征选择算法, 第一个特征子序列的适应度函数是 - 1410630, 第二个特征子序列的适应度函数是 - 1118869, 即第二个特征子序列要优于第一个特征子序列, 这和理论是相吻合的. 而且第二个特征子序列用于测试数据库目标鉴别的总错误数和漏报数都为 0, 而第一个特征子序列用于测试数据库目标鉴别的总错误数为 1, 且全是漏报, 这也进一步证实了第二个特征

表 1 最优特征子序列下 MSTAR 数据的实验结果

训练数据				测试数据			
总错误数	漏报数	总正确率	目标正确率	总错误数	漏报数	总正确率	目标正确率
1	0	0.9975	1	0	0	1	1

表 2 所有特征的鉴别精度

特征编码: $[1\ 1\ 1\ 1\ 1\ 1\ 1\ 1]$ , 适应度值: - 1617723							
训练数据				测试数据			
总错误数	漏报数	总正确率	目标正确率	总错误数	漏报数	总正确率	目标正确率
2	2	0.9949	0.9867	4	4	0.9893	0.9518

表 3 两个特征子序列的鉴别精度

特征子序列编码: $[0\ 1\ 1\ 1\ 1\ 0\ 0\ 1]$ , 本文算法适应度值: - 14.0630, Bhanu 算法适应度值: - 9.7108							
训练数据				测试数据			
总错误数	漏报数	总正确率	目标正确率	总错误数	漏报数	总正确率	目标正确率
2	2	0.9949	0.9867	1	1	0.9973	0.9880
特征子序列编码: $[1\ 1\ 1\ 1\ 1\ 0\ 0\ 0]$ , 文中算法适应度值: - 11.8869, Bhanu 算法适应度值: - 9.7108							
训练数据				测试数据			
总错误数	漏报数	总正确率	目标正确率	总错误数	漏报数	总正确率	目标正确率
2	1	0.9949	0.9933	0	0	1	1

子序列确实优于第一个特征子序列.

推而广之, 不难理解, 如果当种群中存在两个性能较接近的最优个体, 由于它们的特征数和总错误数相同, 因此采用 Bhanu 算法在不同的独立实验中很可能会得到错误的最优特征子序列, 相反, 利用文中特征选择算法却不会. 因此, 文中特征选择算法对于被选特征子序列的优劣评估更全面. 即文中特征选择算法选择的最优特征子序列用于目标鉴别会取得更好的性能.

5 结论

本文针对 SAR 图像自动目标鉴别的应用, 在已有算法的基础上提出了一种基于遗传算法的特征选择方法. 该方法通过综合考虑特征的描述长度、目标鉴别的总错误数以及漏报数设计适应度函数. 理论分析和实测数据的实验结果均表明: 文中特征选择方法对于特征的优劣评价更为全面, 以此选择的最优特征序列应用于目标鉴别时能取得更好的效果.

致谢 作者诚挚地感谢 5 电子学报 6 孙力老师的辛勤工作以及各位审稿老师对于本文负责的审理, 他们的意见使得本文的条理更加清晰, 文字更加凝练, 内容更加准确、充实!

参考文献:

[1]

Dudgeon D E, et al. An overview of automatic target recognition[J]. The Lincoln Laboratory Journal, 1993, 6(1): 3- 10.

[2]

Novak L M, et al. Performance of a high resolution polarimetric SAR automatic target recognition system[J]. The Lincoln Laboratory Journal, 1993, 6(1): 11- 24.

[3]

Novak L M, et al. The automatic target recognition system in SAIP[J]. The Lincoln Laboratory Journal, 1997, 10(2): 187-

202.

- [4] Gao G, et al. Fast detecting and locating groups of targets in high resolution SAR images[J]. Pattern Recognition, 2007, 40(4): 1378- 1384.
- [5] Bhanu B, et al. Genetic algorithm based feature selection for target detection in SAR images[J]. Image and Vision Computing, 2003, 21(7): 591- 608.
- [6] Lin Y Q, et al. Evolutionary feature synthesis for object recognition[J]. IEEE Trans Systems, Man, and Cybernetics Part C: Applications and Reviews, 2005, 35(2): 156- 171.
- [7] 高贵. SAR 图像目标 ROI 自动获取技术研究[D]. 长沙: 国防科学技术大学, 2007.  
Gao Gui. The research on automatic acquirement of target's ROI from SAR imagery[D]. Changsha: National University of Defence Technology, 2007. (in Chinese)
- [8] Kreithen D E, et al. Discriminating targets from clutter[J]. The Lincoln Laboratory Journal, 1993, 6(1): 25- 51.

## 作者简介:



**高 贵** 男, 汉族, 1981 年生于内蒙古集宁, 2002 年 6 月、2003 年 12 月、2007 年 6 月分获国防科技大学工学学士、硕士、博士学位, 国防科技大学信息系统与管理学院讲师, 中国电子学会 / 三遥分会会员。感兴趣的领域包括: SAR ATR、SAR 目标电磁散射建模、遥感信息处理、空间数据挖掘、空间信息系统等。  
E-mail: ggsarxh@sina.com.cn

**周蝶飞** 女, 汉族, 1981 年生, 湖南师范大学硕士研究生。感兴趣领域为遥感信息处理。E-mail: dellarzhou@163.com

**张 军** 女, 汉族, 1975 年生, 国防科技大学信息系统与管理学院副教授。主要研究方向为多媒体信息系统、图像编码、SAR 图像处理等。E-mail: aiya@163.com

**黄纪军** 男, 汉族, 1971 年生, 国防科技大学电子科学与工程学院副教授。主要研究方向为微波测量、目标电磁散射建模与仿真。

**李德仁** 男, 汉族, 1939 年生, 中国科学院院士、中国工程院院士, 武汉大学遥感测绘国家重点实验室教授、博导。主要研究方向为摄影测量与遥感、空间信息系统、地理信息系统、空间数据挖掘等。