

文章编号: 1001-0920(2005)11-1212-04

## 基于GA的遥感图像目标SVM自动识别

郑春红<sup>1</sup>, 焦李成<sup>1</sup>, 郑贵文<sup>2</sup>

(1. 西安电子科技大学 电子工程学院, 西安 710071; 2. 海军驻西安20所军事代表室, 西安 710068)

**摘要:** 为了高效合理地确定支持向量机(SVM)的参数, 使其对复杂的二值遥感图像目标进行自动识别, 采用实值编码遗传算法来实现SVM模型参数的自动选择。与穷举搜索的留一法及随机试凑法相比, 采用遗传算法的SVM模型参数选择更简单、更易于实现, 并使SVM具有更好的推广能力。二值遥感图像目标的分类识别结果表明, 该方法不但可以提高分类识别率, 而且显著地缩短了SVM的训练时间。

**关键词:** 支撑向量机; 遗传算法; 模型选择; 遥感图像; 目标识别

**中图分类号:** TP391      **文献标识码:** A

## Genetic Algorithm-based SVM for Automatic Target Classification of Remote Sensing Images

ZHENG Chun-hong<sup>1</sup>, JIAO Li-cheng<sup>1</sup>, ZHENG Gui-wen<sup>2</sup>

(1. School of Electronic Engineering, Xidian University, Xi'an 710071, China; 2. The Navy Deputy Department Reside in the Twentieth Research Institute, Xi'an 710068, China. Correspondent: ZHENG Chun-hong, E-mail: chzheng@xidian.edu.cn)

**Abstract:** Support vector machine (SVM) has recently been proposed as a new effective learning machine for classification of remote sensing images. However, SVM often requires expensive design phases to choose adequate model parameters to attain high classification accuracy. A real-coded genetic algorithm (RGA) is used to automatically determine the model parameters for SVM, aiming at expediting the model selection process in SVM design with optimal generalization performance. Compared with the commonly used trial-and-error method, the proposed method is easier to implement. Furthermore, the generalization of the RGA-based SVM is much improved. Experimental tests conducted on targets classification of 2-value remote sensing images demonstrate that the proposed approach can conduct automatic model selection with low error while providing significant savings in time.

**Key words:** Support vector machine; Genetic algorithm; Model selection; Remote sensing; Target recognition

### 1 引言

遥感数据分类是遥感技术应用的关键。如何用少量的学习样本在减少分类器训练时间的同时提高分类精度, 一直是遥感图像处理领域亟待解决的问题<sup>[1-3]</sup>。常用的分类方法包括Bayes分类器<sup>[1]</sup>和神经网络<sup>[2]</sup>。Bayes分类器需要建立条件概率密度函数的模型及估计先验概率, 神经网络分类器无须建模, 但需要通过样本学习来确定网络参数, 如果样本选

取不当, 则会导致网络收敛速度慢, 甚至不收敛<sup>[2]</sup>。目前, 支持向量机(SVM)已成为遥感数据分类的一种有效手段<sup>[3]</sup>。

SVM是基于统计学习理论且推广能力良好的一种学习机<sup>[4]</sup>, 它来源于在两类问题中最优越平面具有最好推广能力的思想, 并兼顾了训练误差和泛化能力。SVM在小样本学习及高维特征问题上较传统的分类器具有较大的优势, 因此在模式识别、回归

收稿日期: 2004-11-12; 修回日期: 2005-03-10

基金项目: 国家自然科学基金重点项目(60133010)。

作者简介: 郑春红(1969—), 女, 西安人, 副教授, 博士生, 从事智能信息处理、进化计算的研究; 焦李成(1959—), 陕西白水人, 教授, 博士生导师, 从事智能信息处理、进化计算等研究。

估计及时间序列预测等领域得到了广泛的应用<sup>[3~5]</sup>。

SVM 中模型参数的选择是尚待解决的问题<sup>[6~9]</sup>, 该问题可描述为如何确定合适的参数使获得的 SVM 泛化性能最优。常用的“留一法”(Leave-one-out) 首先需根据人工经验确定一个近似的最优参数集范围, 然后在该参数集上遍历搜索最优参数, 计算量非常庞大<sup>[7]</sup>。Chapelle 等<sup>[6]</sup>应用梯度下降算法实现 SVM 模型参数的自动选择。Bengio<sup>[8]</sup>提出了基于梯度的模型选择判据, 并对线性回归和时间序列预测模型参数进行了优化设计。Keerthi<sup>[9]</sup>采用拟牛顿法进行 Gaussian 核函数 SVM 模型参数选择。上述基于梯度的常规数值方法可能陷于局部最优解, 如果迭代的初值选择不当, 则更难获得令人满意的模型参数<sup>[9]</sup>。

本文提出采用实值遗传算法实现 SVM 模型参数自动选择。基于遗传算法的隐含并行高效性和全局最优的性能选择 SVM 模型参数, 不但大大提高了 SVM 的构造效率, 而且进一步提高了分类器的识别率。二值遥感图像目标自动识别实验验证了所提出方法的有效性和高效性。

## 2 问题提出

### 2.1 支持向量机

分类识别问题可表述为寻找一个分类规则  $I(x)$ , 使它能对未知类别的新样本给出尽可能正确的划分。基于结构风险最小化归纳原则的 SVM, 通过在变换后的高维空间中构造最优分类超平面将给定的样本分开, 构造超平面的依据是两类样本离超平面的距离最大化。

对于给定的含有  $l$  个输入/输出的线性可分样本集  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$ ,  $x_i \in R^n$ ,  $y_i \in \{-1, 1\}$ ,  $i = 1, 2, \dots, l$ 。最优超平面可通过求解下面的凸二次规划最优问题而得到<sup>[4, 6]</sup>:

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2, \\ \text{s.t.} & y_i(w \cdot x + b) = 1 \end{aligned} \quad (1)$$

利用核空间理论, 最优超平面可由下面的对偶规则求得:

$$\begin{aligned} \max W(\alpha) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j), \\ \text{s.t.} & \begin{cases} 0 \leq \alpha_i, i = 1, 2, \dots, l; \\ \sum_{i=1}^l y_i \alpha_i = 0 \end{cases} \end{aligned} \quad (2)$$

相应的分类面方程为

$$f(x) = \text{sgn}\left(\sum_{i=1}^m \alpha_i^* y_i K(x_i, x) + b\right) \quad (3)$$

其中:  $\alpha^*$  为对应  $\alpha = 0$  的向量, 称为支持向量;  $m$  ( $m < l$ ) 为支持向量的数目;  $K(x_i, x) = \phi(x_i) \cdot \phi(x)$  为核函数, 实现输入空间到高维空间的映射;  $b$  为阈值。

上述确定的 SVM 不允许任何训练误差, 因此称为硬间隔 SVM。如果允许分类超平面通过原点, 也就是  $b = 0$ , 那么式(2)中等式约束就不存在, 这样得到的 SVM 称为无阈值硬间隔 SVM。

对于不可分样本, 训练误差必然存在, 这时得到的 SVM 称为软间隔 SVM。它可以看作是硬间隔 SVM 的一个特例, 仍可通过求解式(2)而获得, 惟一不同的是将式(2)中的不等式约束变换为  $0 \leq \alpha_i - C$ , 并将硬间隔 SVM 的核函数进行如下的变换<sup>[6]</sup>:

$$K(x_i, x_j) = K(x_i, x_j) + \frac{1}{C} I \quad (4)$$

其中:  $I$  为单位矩阵,  $C$  为对训练误差的一个恒值惩罚项。

核函数可根据 Mercer's 定理确定<sup>[4]</sup>。通常应用的 Gaussian 核函数可表示为

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5)$$

其中  $\sigma$  为核函数参数。

### 2.2 模型选择性能指标

半径-间隔界反映了留一过程中误差数的上界, 可表示为<sup>[4, 6]</sup>

$$T = \frac{1}{l} \frac{R^2}{\gamma^2} \quad (6)$$

其中:  $R$  为高维特征空间内包含训练数据点的最小球面半径,  $\gamma$  为间隔  $R^2$  的最小化可在

$$\|x_i - x^*\|^2 \leq R^2$$

的约束条件下, 使目标函数

$$R^2 - \sum_{i=1}^l \beta_i (R^2 - (x_i - x^*)^2)$$

最小化来求解,  $x^*$  为待确定的球心位置向量且  $x^* = \sum_{i=1}^l \beta_i x_i$ 。经过推导,  $R^2$  可通过下述的二次优化的最优解来获得<sup>[4, 9]</sup>:

$$\begin{aligned} \max & \sum_{i=1}^l \beta_i K(x_i, x_i) - \sum_{i,j=1}^l \beta_i \beta_j K(x_i, x_j), \\ \text{s.t.} & \begin{cases} 0 \leq \beta_i, i = 1, 2, \dots, l; \\ \sum_{i=1}^l \beta_i = 1 \end{cases} \end{aligned} \quad (7)$$

最小化权重向量等价于最大化间隔, 即

$$\gamma^2 = \|w\|^{-2} \quad (8)$$

所以, 上述的半径-间隔界可重写为

$$T = \frac{1}{l} R^2 \|w\|^2 \quad (9)$$

至此, Gaussian 核函数 SVM 的模型参数选择

可描述为: 如何合理地确定  $\alpha, \beta$  和  $\theta = (C, \sigma)$ , 同时使得  $w$  最大而  $T$  最小, 其中  $\alpha$  和  $\beta$  由标准 SVM 算法求得

### 3 基于实值遗传算法的模型自动选择

遗传算法是一种基于生物进化原理的隐含并行高效随机优化算法 实值遗传算法 (RGA) 的 SVM 模型参数自动选择的流程如图 1 所示, 它包含如下关键步骤:

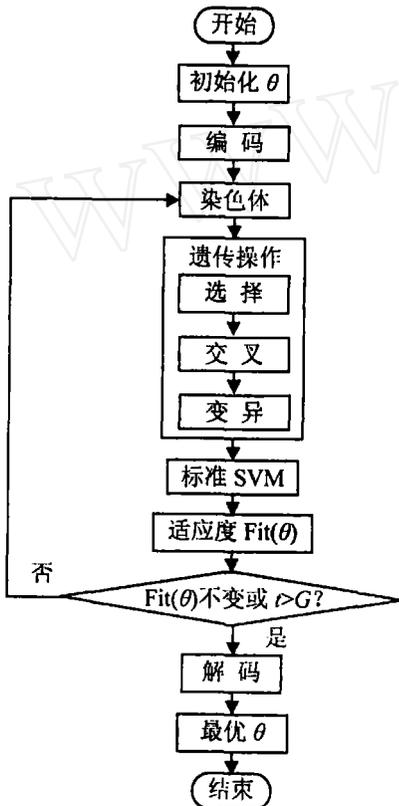


图 1 基于 RGA 的 SVM 模型自动选择流程

Step 1: 编码策略: 采用实值编码策略实现 SVM 模型参数自动选择 根据模型参数  $\theta = (C, \sigma)$  的取值范围, 染色体中的每个基因都由一个 6 位十进制浮点数组成, 代表一个优化参数, 两个基因组合成一条 12 位十进制浮点数构成的染色体

Step 2: 初始种群的产生: 由于采用实值编码, 解空间与染色体空间重合 随机产生初始种群, 并使其均匀分布于解空间

Step 3: 适应度函数: 遗传算法对适应度函数没有连续、可导的要求, 但它是对适应度函数的最大化寻优 因此需将最小化目标函数半径 - 间隔界转化为最大化适应度函数, 如下式所示:

$$Fit(\theta) = \frac{1}{1e^{-6} + T} \quad (10)$$

Step 4: 遗传操作算子: 遗传操作算子是实现遗传寻优的关键, 它主要包括选择算子、交叉算子和变异算子<sup>[10, 11]</sup>.

采用精英算法与适应度比例方法相结合的选择方法作为选择算子.

交叉算子是以随机的方式使染色体的期望性能混合和匹配的主要遗传算子, 它使父代中优秀个体的性状能在子代的新个体中尽可能得到遗传和继承 选择算术交叉算子实现交叉操作<sup>[11]</sup>

$$P = \lambda P_1 + (1 - \lambda) P_2 \quad (11)$$

式中:  $P$  为交叉操作产生的子代个体,  $P_1$  和  $P_2$  分别为两个要进行交叉操作的父代个体,  $\lambda$  为预先定义的  $(0, 1)$  区间内的常数

变异算子使种群中个体的某些基因发生突变, 使遗传算法具有局部的随机搜索能力, 并维持种群的多样性, 避免早熟现象 对父代个体  $P = (g_1, g_2, \dots, g_n)$  中的某个组成元素  $g_j \in [L_j, U_j]$  进行变异操作, 生成新的子代个体  $P = [g_1, \dots, g_j, \dots, g_n]$ , 其变异操作首先等概率地生成二值随机数  $\delta \in \{0, 1\}$ , 然后按下式进行变异<sup>[11]</sup>:

$$g_j = \begin{cases} g_j + \Delta(t, U_j - g_j), & \delta = 0; \\ g_j - \Delta(t, g_j - L_j), & \delta = 1. \end{cases} \quad (12)$$

其中:  $\Delta(t, \vartheta) = \vartheta(1 - \delta^{(1-\tau)/G})^\tau$ ,  $t$  为当前迭代的代数,  $G$  为给定的最大迭代代数,  $\tau$  为预先确定的反映不一致性的参数 同时, 采用如下的自适应变异概率:

$$P_m = \frac{\exp(-t/2)}{\text{pop-size} \sqrt{L}} \quad (13)$$

其中: pop-size 为种群的大小,  $L$  为染色体的长度 使得算法既避免了早熟, 又具有更好的局部搜索能力

Step 5: 终止准则: 选择过去 5 代的平均适应度函数的差异小于  $1e^{-6}$  及最大迭代代数作为终止准则

### 4 实验结果

获取的二值军事目标遥感图像包括完全的目标图像、部分掩盖的目标图像和不同旋转角度的目标图像 图像总数为 310 幅, 其中舰船 150 幅, 飞机 160 幅 图 2 为部分舰船和飞机的二值遥感图像 随机选取 20% 作为训练样本, 剩余 80% 作为验证样本 此外, 另有 513 幅图像作为测试样本

采用双线性插值法将二值图像的分辨率由  $128 \times 128$  降为  $25 \times 25$ , 并进行归一化处理 根据经验, 解空间的范围可确定为  $C \in (0, 0.10000)$ ,  $\sigma \in (0, 0.1000)$ . RGA 的主要控制参数为:  $\lambda = 0.8$ ,  $\delta = 0.3$ ,  $\tau = 5.0$ ,  $L = 12$ , pop-size = 12,  $G = 300$  遗传寻优过程中个体的最优、平均与最差适应度随迭代代数的变化如图 3 所示; 对应于最优适应度函数个体的模型参数如图 4 和 5 所示 RGA 确定的最优模型参数分别为  $C = 7.830 \times 10^{-6}$ ,  $\sigma = 9.56 \times 10^{-2}$

表 1 不同方法确定的 SVM 对样本识别性能比较

方 法	$T$	支持 矢量	模型参数		训练样本集		验证样本集		测试样本集	
			$C$	$\sigma$	识别率/%	错分数	识别率/%	错分数	识别率/%	错分数
试凑法	5.990 8	62	10.0	100.0	51.6	30	51.6	120	87.3	65
	21.532 3	61	100.0	1.0	100.0	0	97.6	6	95.9	21
LOO	20.162 8	61	1.0	1.397 5	100.0	0	97.9	5	96.3	19
RGA	1.864 2	29	7.830 8	9.560 0	100.0	0	98.4	4	98.4	8

应用 RGA、试凑法和 LOO 法所得的对比结果如表 1 所示。为了更充分地说明试凑法的随机性, 表 1 中对于试凑法给出了两种不同的参数组合。

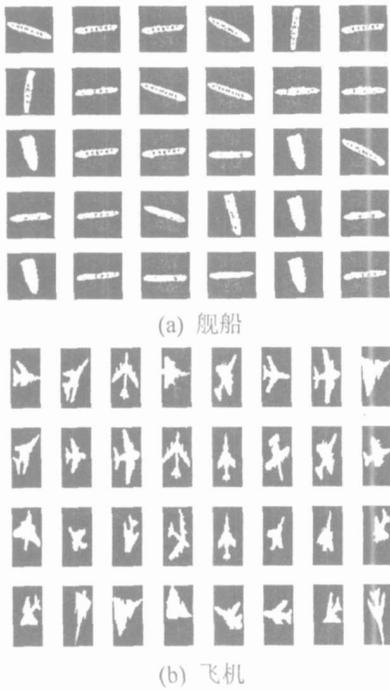


图 2 含有舰船和飞机的遥感图像

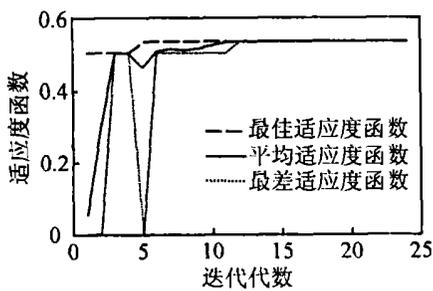


图 3 适应度函数值随迭代代数的变化

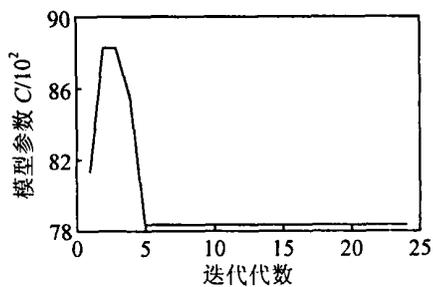


图 4 最大适应度函数对应的模型参数  $C$

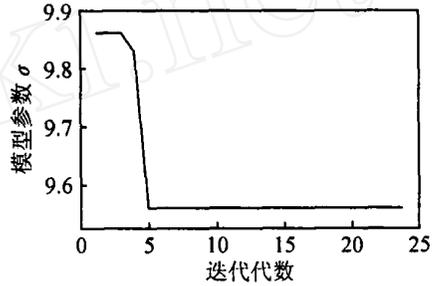


图 5 最大适应度函数对应的模型参数  $\sigma$

从表 1 可以看出, 基于 RGA 的 SVM 具有最小的半径 - 间隔界和最少的支持向量数目。根据 Cristianini 等的研究结果<sup>[5]</sup>, 支持向量数目越少, 支持向量机的泛化性能越好。表 1 中对于验证样本集和测试样本集的实验结果进一步验证了 Cristianini 等的研究结果, 同时也证明了基于 RGA 的 SVM 模型参数的最优性。

此外, 采用试凑法进行一次 SVM 模型参数选择大概需要 1 min 左右, 为了获得较好的性能, 需要反复进行几十甚至上百次的试凑, 这样获得的 SVM 的性能还不一定最优; 而 LOO 法因为采用高效的近似 LOO 误差估计, 使其进行一次模型参数自动选择所需的时间较少, 但从表 1 可以看出, 对于未知的测试样本, 其正确识别率较低; 采用 RGA 的 SVM 参数, 自动确定最优参数约需 15 min, 获得的 SVM 的性能也有较大的提高。

## 5 结 论

为进一步提高二值遥感目标的识别率, 本文提出了 RGA 实现 SVM 模型参数的自动选择方法。采用 RGA 实现 SVM 的模型参数自动选择, 不但确保了 SVM 具有最优的泛化性能, 而且使得 SVM 的训练时间大大缩短。与随机试凑法和 LOO 法的模型参数选择相比, 基于 RGA 的 SVM 在支持向量数、最大间隔、识别率上都体现了一定的优越性。本文提出的方法实现了 SVM 模型参数的自动选择, 为快速设计高性能的实用 SVM 提供了一种简便、高效、通用的方法。二值遥感图像目标的分类识别结果证明了所提出方法的有效性。

(下转第 1220 页)

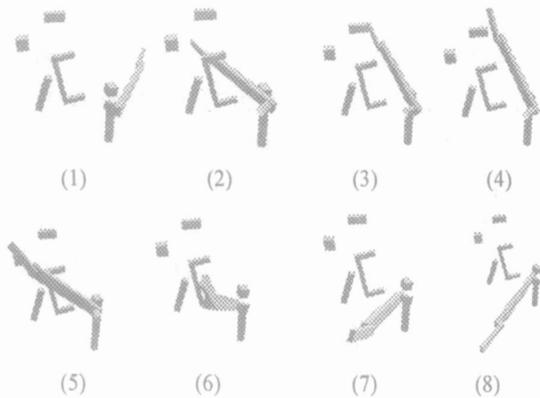


图8 实验结果

#### 4 结 论

本文提出一种新的基于传感信息的实时路径规划方法。由红外线传感器提供机器人手臂周围环境信息,通过计算C-空间内一些方向上的C-空间障碍距离,分阶段控制位姿点到达目标。该方法避免了建立整个位姿点附近的C-空间障碍,适合机器人在未知环境下的实时避障要求。仿真结果证实了该算法具有较好的有效性、实时性和适应性,为工业机器人路径规划提供了新途径。

#### 参考文献(References)

- [1] Belen C, Vidal M, Francisco J B. A General Method for C-space Evaluation and Its Application to Articulated Robots [J]. *IEEE Trans on Robotics and Automation*, 2002, 18(1): 24-31.
- [2] Jiacheng T, Gordon J C. Virtual Environment for Internet-based Robot-II: Path Planning [J]. *Proc of the IEEE*, 2003, 91(3): 389-395.
- [3] Lumelsky V. Sensitive Skin [J]. *IEEE Sensors J*, 2001, 1(1): 41-51.
- [4] Cheung E, Lumelsky V. Motion Planning for A Whole-sensitive Robot Arm Manipulator [A]. *IEEE Int Conf on Robotics and Automation* [C]. Cincinnati, 1990: 344-349.
- [5] Gerke M, Hoyer H. Fuzzy Collision Avoidance for Industrial Robots [A]. *IEEE/RSJ Int Conf on Intelligent Robots and Systems* [C]. Pittsburgh, 1995: 510-517.
- [6] Zavlangas P, Tzafestas S. Industrial Robot Navigation and Obstacle Avoidance Employing Fuzzy Logic [J]. *J of Intelligent and Robotic Systems*, 2000, 27(1/2): 85-97.

(上接第1215页)

#### 参考文献(References)

- [1] Kouskoulas Y, Ulabiy F T, Pierce L E. The Bayesian Hierarchical Classifier and Its Application to Short Vegetation Using Multifrequency Polarimetric SAR [J]. *IEEE Trans on Geoscience Remote Sensing*, 2004, 42(2): 469-477.
- [2] Giacinto G, Roli F, Bruzzone L. Combination of Neural and Statistical Algorithms for Supervised Classification of Remote-sensing Images [J]. *Pattern Recognition Letters*, 2000, 21(5): 385-397.
- [3] Melgani F, Bruzzone L. Classification of Hyperspectral Remote Sensing Images with Support Vector Machines [J]. *IEEE Trans on Geoscience Remote Sensing*, 2004, 42(8): 1778-1790.
- [4] Vapnik V. *The Nature of Statistical Learning Theory* [M]. New York: Springer-Verlag, 1995.
- [5] Cristianini N, Shawe Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* [M]. Cambridge: Cambridge University Press, 2000.
- [6] Chapelle O, Vapnik V, Bousquet O, et al. Choosing Multiple Parameters for Support Vector Machines [J]. *Machine Learning*, 2002, 46(1-3): 131-159.
- [7] Lee J H, Lin C J. *Automatic Model Selection for Support Vector Machines* [R]. Taipei: Taiwan University, 2000.
- [8] Bengio Y. Gradient-based Optimization of Hyperparameters [J]. *Neural Computation*, 2000, 12(8): 1889-1900.
- [9] Keerthi S S. Efficient Tuning of SVM Hyperparameters Using Radius Margin Bound and Iterative Algorithm [J]. *IEEE Trans on Neural Networks*, 2002, 13(5): 1225-1229.
- [10] Holland J H. *Adaptation in Nature and Artificial Systems* [M]. Cambridge: MIT Press, 1992.
- [11] Michalewicz Z. *Genetic Algorithms + Data Structures = Evolutionary Programs* [M]. 3rd ed. Berlin: Springer, 1996.