

克隆选择算法在遥感影像分类中的应用*

张 灵¹, 陈晓宏¹, 翁 毅², 王兆礼³

(1. 中山大学水资源与环境研究中心, 广东 广州 510275;

2. 中山大学地理科学与规划学院, 广东 广州 510275;

3. 华南理工大学土木工程系, 广东 广州 510640)

摘 要: 针对遥感影像分类中的局部极值、鲁棒性等问题, 提出基于克隆选择算法的遥感影像监督分类方法。所提方法将遥感影像各波段亮度值定义为抗原的属性, 通过计算遥感像元与抗体的亲和力识别其类别, 并采用实数制方式对抗体进行变异, 在保证亲和力上升的同时, 也保证了系统的多样性。该算法应用于广州市遥感影像数据分类的实验结果显示: 在抗体的进化过程中, 抗体的亲和力和识别能力不断提高, 最终的分类精度达到92.9%。与最大似然法相比, 克隆选择算法的分类精度更高。

关键词: 遥感; 图像识别; 人工智能; 克隆选择

中图分类号: P283.8 **文献标识码:** A **文章编号:** 0529-6579 (2008) 03-0128-05

随着人类对自然资源开发利用的深入, 土地利用、覆盖情况发生了快速的变化, 新的地表信息的获取日益受到关注。卫星遥感数据为人们提供了一个新的信息渠道。但由于自然环境的复杂性, 从遥感数据中提取的地表信息总是存在一定的不确定性^[1]。如何提高遥感图像分类的精度, 以使从光谱信号中提取的信息准确客观地反映实际情况, 是学术界的热点问题。

图像分类是指对图像信息进行处理、分辨、检测, 确定图像上物体的属性或特征, 对图像进行分类或识别^[2]。传统的遥感图像分类方法主要分为监督分类和非监督分类, 非监督分类虽然不需要分析者十分了解所分类的区域, 但是“同物异谱、异物同谱”问题限制了所产生的光谱集群组的精度^[3]; 而监督分类虽然利用人的先验知识确定聚类中心, 分类精度有所提高, 但是由于未考虑训练数据全局特征, 容易造成局部最优的情况^[4]。目前最常用的遥感图像分类方法是基于参数化密度分布函数判别的最大似然法, 但它要求图像数据呈正态分布^[5], 而且由于没有地学知识的支持, 有时难以真实反映一些特殊的地学分布^[6]。

目前, 基于多智能体的一些人工智能算法引起了地理学者的兴趣, 包括蚁群智能^[7]、粒子群^[5]以及人工免疫系统^[8], 并已逐步被引入到空间数据挖掘中, 如遥感分类和地理规则的自动提取。人

工免疫系统 (Artificial Immune System, AIS) 是模仿自然免疫系统功能的一种人工智能方法, 它通过实现外界物质的自然防御机理的学习技术, 提供了噪声忍耐、自学习、自组织、记忆等进化学习机理, 结合了分类器、神经网络和机器推理等系统的优点^[9-10]。与传统的分类方法不同, 人工免疫算法是一种非参数的分类方法, 并不要求训练数据为正态分布, 因此训练样本的选择和参数估计不会直接影响分类结果。目前 AIS 已在故障诊断^[11]、模式识别^[12]、数据挖掘^[13]等领域得到了成功的应用, 但将 AIS 用于遥感图像分类的研究尚不多见。钟燕飞等^[14]提出了基于 AIS 的遥感影像非监督分类方法。本文尝试引入人工免疫系统的克隆选择算法并设计一种监督分类方法, 以期解决遥感影像分类中的先验性、局部极值、鲁棒性等实际难点。

1 人工免疫算法

1.1 基本架构

免疫系统抵御外部入侵使其机体免受病原侵害的应答反应称为免疫。免疫系统中的淋巴细胞主要有 B 细胞和 T 细胞^[15]。B 细胞的主要功能是产生抗体, T 细胞则直接对抗原实施攻击。T 细胞在体内循环, 一旦检测出任何非己物质就执行免疫功能, 以保护机体不受抗原的侵害。免疫算法大多将 T 细胞、B 细胞、抗体等功能合而为一, 统一抽象

* 收稿日期: 2007-11-12

基金项目: 国家自然科学基金资助项目 (50579078); 广东省自然科学基金资助项目 (04009805)

作者简介: 张灵 (1974 年生), 男, 博士研究生; 通讯联系人: 翁毅; E-mail: Onlyzsu@yahoo.com.cn

出检测器概念, 主要模拟生物免疫系统中有抗原处理的核心思想, 包括抗体产生、自体耐受、克隆选择、免疫记忆等^[15]。一般的人工免疫算法架构如图1所示。

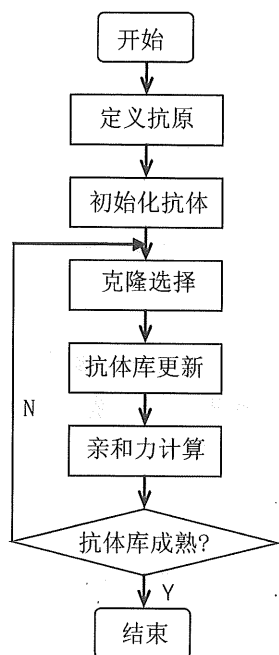


图1 人工免疫系统构架

Fig. 1 Structure of Artificial Immune System

1.2 克隆选择原理

克隆选择原理由 Jerne 于 1955 年提出, 之后经 Burnet 发展和完善^[16], 是免疫系统的基本理论之一。克隆选择的基本思想是当任何一种抗原侵入时, 机体都能在内部选择出能识别和消灭相应抗原的免疫细胞克隆, 使之激活、分化和增殖, 进行免疫应答以最终清除抗原^[17]。克隆选择描述了获得性免疫的基本特性: 只有成功识别抗原的免疫细胞才得以增殖, 经历变异后的免疫细胞分化为浆细胞(抗体效应细胞)和记忆细胞两种。克隆选择原理可通过采用变异等遗传算子和相应的群体控制机制实现^[18]。

1.3 克隆选择算法

算法的设计是整个人工免疫系统的重要组成部分, 对完成系统的进化以及特定问题的收敛有着至关重要的作用。人工免疫系统的实现算法主要有阴性选择算法、克隆选择算法、人工免疫网络算法等。其中 De Castro 提出的克隆选择算法的核心是比例复制和比例变异算子^[19]。克隆选择算法的具体实现过程可参考文献 [19]。

2 基于克隆选择的遥感图像分类

2.1 抗原定义

将需要解决的问题抽象成符合免疫系统处理的抗原形式, 抗原识别则对应为问题的求解。遥感图像的分类是依据每个像元各个不同波段的光谱亮度和空间结构特征, 按照一定的规则和算法将像元划分为不同类别。本文使用影响遥感图像分类的经过预处理的各波段亮度值定义抗原, 可表示为:

$$Ag = (b_1(Ag), b_2(Ag), \dots, b_N(Ag): C) \quad (1)$$

式(1)中, $(b_1(Ag), b_2(Ag), \dots, b_N(Ag))$ 为抗原 Ag 所对应的遥感图像各个波段同一点亮度值。C 为其抗原在分类所对应的类别。本文对亮度值进行了归一化处理。

2.2 抗体库的初始化

采用随机的方法初始化抗体群体, 抗体初始化公式如下:

$$Ab = (b_1(R), b_2(R), \dots, b_N(R): S) \quad (2)$$

式(2)中, $b_1(R), b_2(R), \dots, b_N(R)$ 分别为 0 ~ 1 之间的随机数。S 为保留值, 在抗体库成熟时保留与其亲和力最高的抗原对应的类别。

2.3 亲和力的定义

在人工免疫系统中, 抗原的识别效果由抗体与抗原之间的匹配程度来描述, 这种匹配程度称为亲和力, 亲和力较高的抗体有可能被选择保留及增殖。本文采用下式计算抗体与抗原之间的亲和力:

$$Af(Ag, Ab) = \frac{1}{1 + \text{dis}(Ag, Ab)} \quad (3)$$

$$\text{其中, } \text{dis}(Ag, Ab) = \sqrt{\sum_{n=1}^N (b_n(Ag) - b_n(R))^2} \quad (4)$$

2.4 克隆选择

从初始化的抗体中选取亲和力较高的组成抗体库, 再对抗体库中的抗体进行克隆选择。分别对抗体库中的每个抗体进行多次克隆, 并根据亲和力的高低发生变异, 亲和力越高, 变异的可能性越小。之后从变异个体中选取亲和力较高的抗体保留在抗体库中。如此进行多次克隆选择直到抗体库成熟。本文以对测试数据的整体亲和力作为判断抗体库是否成熟的标志。本文实例中每次对每个抗体均克隆出 30 个新抗体进行变异。为了保证系统的鲁棒性, 避免产生过度集聚的现象, 本文采用随机性较大的实数变异机制, 变异规则为:

$$b_i(R) = \begin{cases} b_i(R-1), Af_{gh}/1.5 > \text{Rand}() \\ \text{Rand}(), Af_{gh}/1.5 < \text{Rand}() \end{cases} \quad (5)$$

变异完成之后分别计算各个新克隆抗体的亲和力,

保留亲和力高的进入抗体库。由于采用筛选较优新抗体的选择方法,系统不会因为变异的随机性而导致亲和力下降,从而在保证系统多样性的同时,也使得整体亲和力不断上升。

2.5 遥感图像分类

应用上述经过克隆选择的成熟抗体库进行遥感分类。将每个遥感的像元表示为与抗原类似的形式: $Ag = (b_1(Ag), b_2(Ag), \dots, b_N(Ag))$, 与成熟的抗体库中的每个抗体进行亲和力的计算,与此像元亲和力最高的抗体的类别即视为此像元的分类类别。

3 克隆选择遥感分类应用实例

选取经过预处理的广州市海珠区 2005 年 7 月 18 日 TM 卫星遥感影像数据进行实验 (图 2)。

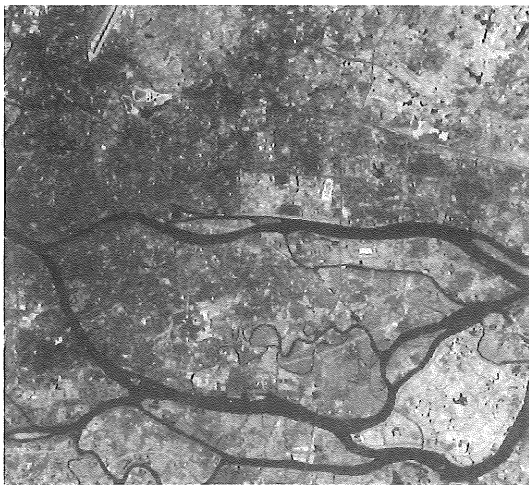


图 2 广州市 TM 影像数据 (5, 4, 3)
Fig. 2 Guangzhou TM image (5, 4, 3)

研究区内的主要地物类型为城镇、珠江及人工湖等水体和部分农田、果园,因此将广州的土地利用类型划分为城市用地、山体、水体、果园、开发用地和农田等 6 类。选择除第 6 波段以外的 6 个波段作为抗原的空间向量。对 TM 卫星图像随机采样,加上人工识别的采样点的分类结果,则抗原 $Ag = (band_1, band_2, \dots, band_7; C)$, C 为对应的分类。采样之后将各个抗原进行归一化处理。训练数据的样本为 1500 个,验证数据集的样本为 1200 个。

按式 (2) 随机生成初始抗体,选择亲和力较高的组成抗体库,进行克隆更新。本文进行了 100 次抗体库的克隆选择,每次克隆后都会计算测试数据的精度,精度随着不断的克隆变异呈波动上升,抗体和抗原的亲和力呈上升的趋势,亲和力的变化

曲线如图 3 所示。

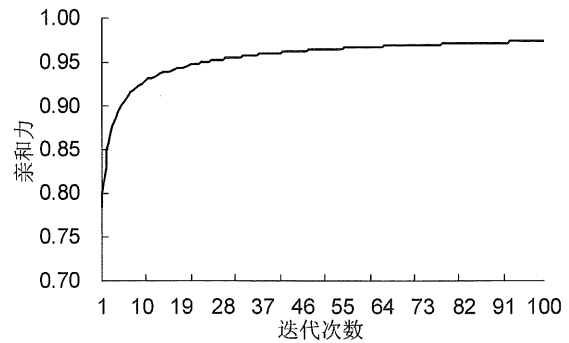


图 3 亲和力变化曲线

Fig. 3 The trend of affinity

从图 3 可以看出,随着克隆选择的进行,抗体库的整体亲和力逐渐上升,直至达到平缓的状态。此时,抗体库已经成熟,即可对原始图像进行分类。最终城市用地、山体、水体、果园、开发用地和农田等 6 类地物的总体分类精度为 92.9%,分类结果见图 4 (a)。为便于对比,选取相同的训练数据,采用应用广泛的极大似然法对同一幅 TM 影像进行了分类,分类结果见图 4 (b)。按先前选取的验证数据对两种方法的分类结果进行检验,检验结果见表 1。

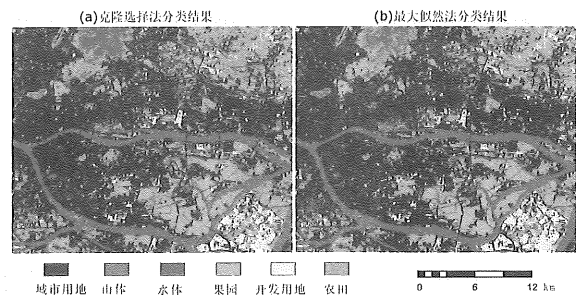


图 4 克隆选择和最大似然法分类结果比较

Fig. 4 Comparison of clonal selection algorithm and maximum likelihood classification

从图 4 和表 1 可以看出,采用 AIS 克隆选择算法进行遥感影像分类时,城市用地、开发用地、山体和山体等大面积连续区域用地类型的分类基本准确,分类精度均比最大似然法高。尤其克隆选择算法对公路与桥梁比较敏感,能将这些较细的线状城市用地识别出来,很大程度上提高了 TM 影像的分类精度。不足的是,尽管果园和农田的分类精度超过了 80%,达到分类的基本精度要求,但是与其他用地类型相比,精度仍较低,不能完全准确地反映果园和农田的实际用地情况。

表1 克隆选择和最大似然法遥感影像分类精度

Tab. 1 The precision of TM image classification

%

用地类型	城市用地	山体	水体	果园	开发用地	农田	总体精度
克隆选择算法	94.5	95.3	98.2	83.1	97.7	82.8	92.9
最大似然法	92.4	87.5	96.2	82.7	97.6	84.6	89.6

总体来说, 基于克隆选择算法的遥感影像分类的总精度为 92.9%, 比最大似然法高出 3.3 个百分点。而从各项用地的分类精度来看, 克隆选择算法比较适合应用于快速发展的中心城市的遥感影像分类。

4 结 语

人工免疫系统的克隆选择算法可解决遥感图像分类中的局部极值、鲁棒性等问题, 同时人工免疫算法不要求训练数据为正态分布, 因此训练样本的选择和参数估计不会直接影响分类结果。将基于克隆选择的遥感图像分类算法应用到广州市 TM 遥感影像数据分类的实例说明, 本文设计的克隆选择算法良好地运用了 AIS 的学习认知的计算智能, 在保证系统多样性的同时使整体亲和力和维持逐渐上升的趋势, 获得了较高的分类精度。通过与最大似然法的比较, 克隆选择算法采用了与传统方法寻找聚类中心不同的思想, 克服了实际样本数据可能与正态分布偏离较大的问题, 保证了对训练数据的全局把握, 从而避免了图像的错分, 获得了更高的分类精度, 适合应用于快速发展的中心城市。但克隆选择算法的运算效率以及对果园和农田等用地类型的分类精度还有待进一步提高。

致谢: 感谢刘小平博士给予本文作者的大量帮助。

参考文献:

- [1] 李德仁, 关泽群. 空间信息系统的集成与实现[M]. 武汉: 武汉测绘科技大学出版社, 2000.
- [2] 陈述彭, 赵英时. 遥感地学分析[M]. 北京: 测绘出版社, 1990.
- [3] 术洪磊, 毛赞猷. GIS 辅助下的基于知识的遥感影像分类方法研究[J]. 测绘学报, 1997, 26(4): 328-335.
SHU Honglei, MAO Zanzhi. Knowledge based image classification approach supported by a GIS [J]. Acta Geodaetica et Cartographica Sinica, 1997, 26(4): 328-335.
- [4] ATKINSON P M, LEWIS P. Geostatistical classification for remote sensing: an introduction [J]. Computers & Geosciences, 2000, 26: 361-371.
- [5] 刘小平, 黎夏, 彭晓晴, 等. 一种基于生物群集智能优化的遥感分类方法[J]. 中国科学: D 辑 地球科学, 2007, 37(10): 1400-1408.
- [6] 骆剑承, 周成虎, 杨艳. 遥感地学智能图解模型支持下的土地覆盖/土地利用分类[J]. 自然资源学报, 2001, 16(2): 179-183.
- [7] 刘小平, 黎夏, 叶嘉安, 等. 利用蚁群智能挖掘地理元胞自动机的转换规则[J]. 中国科学 D 辑: 地球科学, 2007, 37(6): 824-834.
- [8] 钟燕飞, 张良培, 龚健雅, 等. 基于人工免疫系统的遥感图像分类[J]. 遥感学报, 2005, 9(4): 374-380.
ZHONG Yanfei, ZHANG Liangpei, GONG Jianya, et al. Remote sensing image classification based on artificial immune system [J]. Journal of Remote Sensing, 2005, 9(4): 374-380.
- [9] FORREST S, PERELSON A S, ALLEN L, et al. Self-nonsel self discrimination in a computer [C]// Proceedings of IEEE Symposium on Research in Security and Privacy. Oakland, USA, 1994: 202-212.
- [10] DASGUPTA D. Artificial Immune Systems and Their Applications [M]. Berlin: Springer-Verlag, 1999.
- [11] 熊浩, 孙才新, 陈伟根, 等. 电力变压器故障诊断的人工免疫网络分类算法[J]. 电力系统自动化, 2006, 30(6): 57-60.
XIONG Hao, SUN Caixin, CHEN Weigen, et al. Artificial Immune Network Classification Algorithm for Fault Diagnosis of Power Transformers [J]. Automation of Electric Power Systems, 2006, 30(6): 57-60.
- [12] CARTER J H. The Immune system as a model for pattern recognition and classification [J]. Journal of the American Medical Informatics Association, 2000, 7(3): 28-41.
- [13] TIMMIS J, NEAL M, HUNT J. An artificial immune system for data analysis [J]. Biosystem, 2000, 55:

- 143 - 150.
- [14] ZHONG Y F, ZHANG L P, HUANG B, et al. An unsupervised artificial immune classifier for multi hyperspectral remote sensing imagery [J], IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(2): 420 - 431.
- [15] 李涛. 计算机免疫学[M]. 北京: 电子工业出版社, 2004.
- [16] BURNET F M. The Clonal Selection Theory of Acquired Immunity [M]. UK: Cambridge University Press, 1959.
- [17] 林学颜, 张玲. 现代细胞与分子免疫学[M]. 北京: 科学出版社, 1999.
- [18] 肖人彬, 王磊. 人工免疫系统: 原理、模型、分析及展望[J]. 计算机学报, 2002, 25(12): 1281 - 1293.
- XIAO R B, WANG L. Artificial immune system: principle, models, analysis and perspectives [J]. Chinese Journal of Computers, 2002, 25(12): 1281 - 1293.
- [19] DE CASTRO L N, VON ZUBEN F J. The clonal selection algorithm with engineering applications [C]// Genetic and evolutionary computation conference. Nevada, USA, 2000: 36 - 37.

Clonal Selection Algorithm for Classification of Remote Sensing Imagery

ZHANG Ling¹, CHEN Xiao-hong¹, WENG Yi², WANG Zhao-li³

(1. Center for Water Resources and Environment Research;

2. School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China;

3. Department of Civil Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: A novel supervised classifier for remote sensing data by employing clonal selection algorithm (CSA) is presented for solving the problems, e. g. local optimum and robustness, in remote sensing imagery classification. In the classifier, band brightness is defined as antigen's attribute, and the image is classified into class with the maximum affinity by calculating the affinity between remote sensing pixel and antibody. Antibody's real-encoding mutation integrates affinity's ascending and system's variety. The concerned experiment shows that antibody's affinity increases along with its evolution, and CSA classification's overall accuracy is 92.9%. Comparing with conventional Maximum Likelihood, CSA can get better precision.

Key words: remote sensing; pattern recognition; artificial intelligence; clonal selection algorithm