

基于人工免疫系统的遥感图像检索

耿利川^{1), 2)} 吴云东^{1), 3)} 耿则勋¹⁾ 王永刚¹⁾

¹⁾ (信息工程大学测绘学院, 郑州 450052) ²⁾ (66240 部队, 北京 100042) ³⁾ (集美大学理学院, 厦门 361021)

摘要 为了提高遥感图像检索的有效性和准确性, 提出了一种基于人工免疫系统的遥感图像检索算法。该算法根据相关反馈技术及免疫机理, 利用克隆选择算法对用户反馈的图像特征进行泛化学习, 从而提高了系统对用户语义的理解能力。实验结果表明, 该算法不仅能有效理解用户的反馈信息, 而且能提高检索的准确性。

关键词 人工免疫系统 遥感图像检索 相关反馈 克隆选择

中图法分类号: TP391.3 TP751.1 文献标志码: A 文章编号: 1006-8961(2010)01-0155-06

Remote Sensing Image Retrieval Based on Artificial Immune System

GENG Li-chuan^{1), 2)}, WU Yun-dong^{1), 3)}, GENG Ze-xun¹⁾, WANG Yong-gang¹⁾

¹⁾ (Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052)

²⁾ (66240 Troops Beijing 100042) ³⁾ (College of Sciences, Jimei University, Xiamen 361021)

Abstract A novel remote sensing image retrieval method based on artificial immune system (AIS) is proposed in this paper. Based on the relevance feedback technology and the immune mechanism, clonal selection algorithm is used to learn and memorize the user feedback image feature, the recognition of customers' semantic target for system is improved. Experimental results show that this approach can recognize the user feedback information efficiently and improve the retrieval accuracy.

Keywords artificial immune system, remote sensing image retrieval, relevance feedback, clonal selection

0 引言

由于图像的底层特征和用户的抽象主观需求之间存在语义鸿沟, 导致了现有的基于内容的图像检索技术很难达到实用的要求。为了弥补底层特征和高层语义之间的差距, Rui Yong 最早将相关反馈技术 (relevance feedback) 系统地引入到了图像检索中, 其目的是利用人工交互, 让用户进一步提供当前检索需求的视觉模式信息, 以便使系统能够根据用户的选择来判断用户的特征关注类型^[1]。但是目前这些相关反馈方法普遍认为, 由于用户的检索目标围绕一个中心点聚集成簇, 因此可以通过移动检

索点和修正距离度量函数来改进检索效果。事实上由于受多方面因素的影响, 目前使用的图像底层特征还难以全面地刻画用户的检索目标, 因此满足用户需求的图像在特征空间中的分布是非常不均匀的, 往往出现围绕多个中心^[2]的现象。

人工免疫系统 (artificial immune system, AIS), 是模仿自然免疫系统功能的一种智能方法, 它是一种受生物免疫系统启发, 通过学习外界物质的自然防御机理的学习技术, 由于该方法能提供噪声忍耐、无教师学习、自组织、记忆等进化学习机理, 以及能结合分类器、神经网络和机器推理等系统的一些优点, 因此提供了一种新颖的解决问题的方法和途径。其研究成果已应用于设计控制、数据处理、优化学习

收稿日期: 2008-06-27 改回日期: 2009-02-18

第一作者简介: 耿利川 (1982—), 男, 信息工程大学测绘学院摄影测量与遥感专业硕士研究生。主要研究方向为数字摄影测量、图像检索。E-mail: glch1982@163.com

和故障诊断等许多领域^[3]。

文献 [4-6]将人工免疫系统用于互联网中的信息检索,即利用人工免疫网络对查询网络进行建模。文献 [7]提出了一种基于免疫学架构的文本信息检索模型,而且这些文献研究内容都是基于关键字的文本检索,没有涉及基于内容的图像检索。文献 [2]提出了一种基于免疫算法的图像检索模型并实现了一个原型化检索模型,但由于该方法需要利用模糊聚类方法对图像库进行聚类,因此检索的准确性很大程度上取决于聚类的精度,而且由于需要用户反馈的次数相对比较多,致使算法精度收敛比较慢。除此之外,目前还未见将人工免疫系统应用在遥感图像检索中的报道。

鉴于人工免疫系统在无教师学习、自动识别、记忆中表现出的优越性,本文提出了一种利用克隆选择算法对用户反馈的图像信息进行泛化学习、记忆,以实现查找最优图像的遥感图像检索算法。实验表明,本算法只需较少次数的用户反馈,即可得到较高精度的检索结果。

1 克隆选择理论及算法

1.1 克隆选择理论

1957年,奥地利免疫学家 Bumer 在免疫生物学研究及遗传学研究的基础上,提出了抗体生成的克隆选择学说^[8]。此学说不仅阐明了抗体生成的生物学机制,而且对许多重要的免疫生物学现象,如抗原识别、免疫记忆、自身耐受、自身免疫等一系列问题都做出了解答,从而奠定了现代免疫学的理论基础。

克隆选择学说的中心思想是,抗体是天然产物,它是以受体的形式存在于细胞表面,抗原可与之进行选择性的反应。抗原与相应抗体受体的反应可导致细胞克隆性增殖,而且该群体具有相同的抗体特异性,其中某些细胞克隆分化为抗体生成细胞,而另一些则形成免疫记忆细胞,以便参加之后的二次免疫反应。

1.2 克隆选择算法

克隆选择算法最初用于解决机器学习、模式识别和分类问题,后来在解决优化问题方面也得到广泛的使用。所有的基于克隆选择理论的人工免疫系统都是基于种群的,而每个种群即代表一个候选解。克隆选择算法的过程大致为先使用克隆算子来产生

高亲和力的个体,再利用高频变异算子进行变异,最后选择改进的成熟后代。

De Castro 和 Von Zuben 根据克隆选择理论提出了克隆选择算法 (CLONALG)^[9-10]。

克隆选择算法流程如下 (图 1):

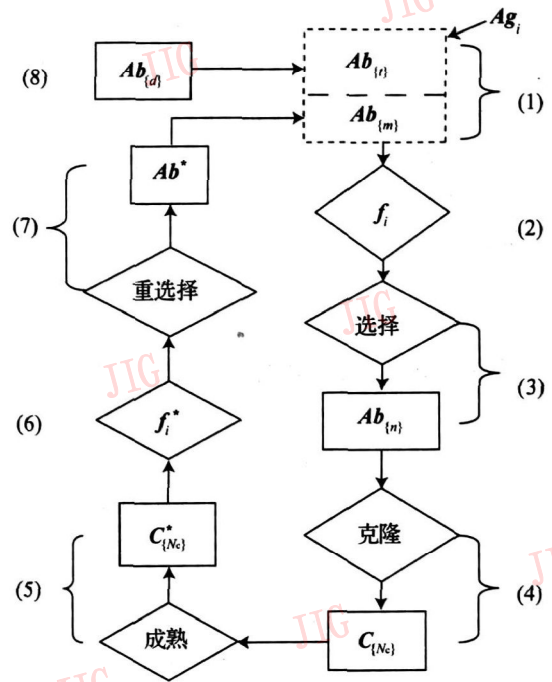


图 1 克隆选择算法流程图

Fig 1 Block diagram of the clonal selection algorithm

1) 随机选择某个抗原 $Ag_i (Ag_i \in Ag_{(M)})$, 其中 $Ag_{(M)}$ 为抗原集合, M 为抗原总数, 而抗体集合为 $Ab = Ab_{(t)} + Ab_{(m)} (t + m = I)$ 。

2) 计算亲和力矢量 f_i , 包括抗原 Ag_i 与抗体集合 Ab 中所有抗体的亲和力。

3) 从抗体集合 Ab 中选择 n 个与抗原 Ag_i 的亲和力最高的抗体, 产生一个新的抗体集合 $Ab_{(n)}$ 。

4) 对 n 个被选择的抗体进行克隆, 其亲和力高的抗体克隆的数量高, 通过克隆即可得到克隆集合 $C_{[j]}$ 。将 n 个亲和力高的抗体按照升序排列, 其克隆的数量由下式确定:

$$N_c = \sum_{i=1}^n \text{round} \left| \frac{\beta \cdot f_i}{f_i} \right| \quad (1)$$

其中, N_c (下角 c 代表 clonal) 是克隆的总数量, β 是增殖系数, I 是抗体的数量, $\text{round}(\cdot)$ 为取整操作。

5) 集合 $C_{[N_c]}$ 变异后产生亲和力成熟的抗体集合 $C_{[N_c]}^*$, 变异过程中, 亲和力大的抗体变异机会小。

6) 计算亲和力矢量 f_i^* , 它是所有成熟抗体 $C_{[j]}^*$ 与抗原 Ag_i 的亲和力。

7) 从 f_i^* 中选择最高亲和力的抗体 Ab_j^* 作为进入记忆抗体集合 $Ab_{(m)}$ 的候选抗体。如果这个候选抗体的亲和力大于记忆抗体集合中的抗体, 则替换这个记忆抗体。

8) 最后, 随机生成 d 个抗体 $Ab_{(d)}$, 替换 $Ab_{(m)}$ 中的 d 个最低亲和力的抗体。

2 基于克隆选择算法的遥感图像检索

2.1 基本原理

本文算法是将检索图像数据库视为抗原, 将目标图像视为抗体, 检索过程就是免疫系统不断优化抗体、识别抗原, 直到产生可以捕获抗原的最佳抗体的过程。

1) 抗原捕获 本文算法假定由用户提供一个目标图像, 从中提取特征后, 即生成 I 个抗体 ab_i ($i = 1, 2, 3, \dots, I$), 图像数据库中有 M 个对象 X_j ($j = 1, 2, 3, \dots, M$), 对这 M 个对象提取特征后, 生成 M 个 k 维特征向量 ag_j ($j = 1, 2, 3, \dots, M$), 作为抗原, 检索过程就是由这 I 个抗体来捕获 M 个抗原。选择欧氏距离为抗体和抗原的亲和力指标, 按抗原与这 I 个抗体中最大亲和力的大小, 由大到小排序, 显示前 L 个结果。亲和力的计算公式如下:

$$f_{ij} = 1 - |ab_i - ag_j| \quad (2)$$

若抗原和抗体都经过归一化, 则亲和力值介于 0 和 1 之间。

2) 抗体优化 这里借鉴 De Castro 和 Von Zuben 的克隆选择模型, 利用用户反馈图像的特征信息, 对抗体进行克隆选择, 并进化。

重复上面的抗原捕获及抗体优化过程, 直到用户检索到满意结果。

2.2 基于小波变换的特征提取

本文利用小波变换对图像进行 3 级小波分解, 子带 $B(LL, LH, HL, HH)$ 中的第 mn (m 为行号, n 为列号) 个小波系数, 分解级数为 l 时记为 $(c_{m,n}^{(l)}, d_{m,n}^{(1)}, d_{m,n}^{(2)}, d_{m,n}^{(3)})$, 低频系数的能量记为 $e_{m,n}^{(l)} = c_{m,n}^{(l)} \cdot c_{m,n}^{(l)}$; 然后提取低频系数的一阶矩、二阶矩、信息熵, 其中能量的一阶矩、二阶矩、信息熵为特征。数学上的前二阶矩定义如下:

$$U_i = 1/N \cdot \sum_{j=1}^N P_{ij} \quad (3)$$

$$Z_i = [1/N \cdot \sum_{j=1}^N (P_{ij} - U_i)^2]^{1/2} \quad (4)$$

本文中的信息熵计算如下:

$$S_i = 1/N \cdot \sum_{j=1}^N (-P_{ij} \cdot \ln(P_{ij})) \quad (5)$$

式中, P_{ij} 为小波变换后的低频系数, 提取特征后, 每幅彩色图像生成 18 维特征向量, 记为

$$v_i = (U_i^{1r}, Z_i^{1r}, S_i^{1r}, U_i^{2r}, Z_i^{2r}, S_i^{2r}, U_i^{1g}, Z_i^{1g}, S_i^{1g}, U_i^{2g}, Z_i^{2g}, S_i^{2g}, U_i^{1b}, Z_i^{1b}, S_i^{1b}, U_i^{2b}, Z_i^{2b}, S_i^{2b}) \quad (6)$$

其中, $i = 0, 1, \dots, M-1, M$ 为图像数。

2.3 图像特征向量归一化^[12]

用于图像相似性测度的各特征值均需归一化到同一值域区间内, 这是因为如果某一特征的值域范围较大, 那么它就会削弱其他值域范围较小的特征的贡献, 从而使得通过各个特征相似性值进行线性组合来计算总相似性值变得没有意义。本文采用简单易行的极值归一化方法来进行特征向量归一化。

设数据库中图像数为 M , 特征向量的维数为 K , 第 i 个图像的第 k 个特征值为 v_{ik} , 则可将向量中的每个元素用式 (7) 归一化到区间 $[0, 1]$ 内,

$$v_{ik} = \frac{v_{ik} - \min(V_k)}{\max(V_k) - \min(V_k)} \quad (7)$$

$$i = 0, 1, \dots, M-1; k = 0, 1, \dots, K-1$$

其中, $\min(V_k)$, $\max(V_k)$ 分别为由所有图像的第 k 维特征组成的向量 V_k 的最小值和最大值。

为了将各特征向量之间的欧氏距离归化到区间 $[0, 1]$ 内, 需将归一化后的特征向量除以特征向量模的 $\sqrt{2}$ 倍, 如下式所示:

$$v_i = \frac{v_i}{\sqrt{2} \cdot |v_i|} \quad (8)$$

其中, v_i 为第 i 个图像的特征向量, $|v_i|$ 为向量 v_i 的模。

2.4 算法实现

1) 输入 M 个抗原 算法假定图像数据库中的 M 个特征为 M 个被捕获的抗原输入 $Ag = \{ag_1, ag_2, \dots, ag_M\}$;

2) 产生初始 I 个抗体 将目标图像均分为 I 块, 分别提取特征, 生成抗体库 $Ab = \{ab_1, ab_2, \dots, ab_I\}$;

3) 识别抗原 分别计算抗原 ag_i 与抗体库 Ab 的亲和力, 按抗原与抗体库中抗体亲和力的最大值, 从大到小, 依次排序显示前 L 个结果;

4) 用户反馈 用户根据显示结果从中指定相

关的反馈对象;

5)抗体选择 由用户反馈对象特征,并计算与现有抗体库 Ab 的亲合力,亲合力按从高到低排序,然后从中选取具有最大亲和力的个体 ab_{max} ;

6)抗体抑制 为提高整个种群的整体亲合力,以一定的抑制率 r 删除种群中小亲和力的个体,并将用户的反馈特征加入抗体库;

7)抗体克隆及变异 将亲合力最高的个体按一定的克隆率 c 进行克隆,取代被删除的亲合力小的个体,对克隆的个体进行高斯变异,变异率为 p ,变异后的个体其亲合力如果高于原抗体,则取代原抗体进入抗体群;

8)利用生成的高亲和力的特征组,重新检索,如此循环,直到用户结束检索。

本文算法流程示意如图 2 所示。

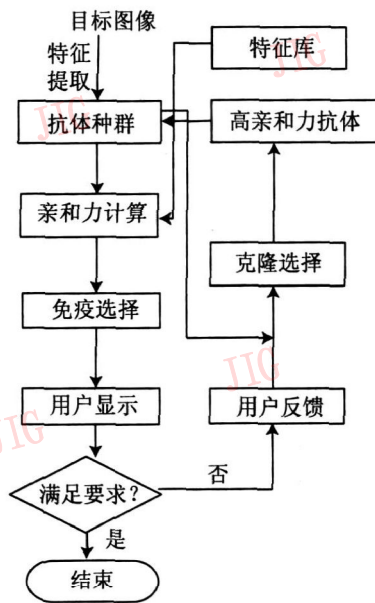


图 2 本文算法流程图

Fig 2 Block diagram of our algorithm

3 实验结果

本文使用的实验图像库是在 2005-05-08 用无人飞机搭载 Canon 1D Mark-II 相机,在驻马店拍摄的航空遥感影像,从中抽取 1 000 幅 512×512 pixels 大小的子图组成的,图像中包含街区、道路和植被 3 类区域。本检索实验在计算机主频 2.93GHz、内存 512M、Windows XP、matlab 7.0 环境中实现。

3.1 遥感图像检索

首先将目标图像等分为 9 块,然后利用 db4 小波对目标图像整体及 9 个子块进行 3 层小波变换,并提取低频子带系数特征向量,组成 10 个特征向量。

实验中的参数设置如下:克隆率 c 为 20%,抑制率 r 为 30%,变异率 p 为 10%。

实验时,从每种地类图像中选取 10 幅有代表性的图像进行检索,并显示前 25 个检索结果,经过 7 次反馈。每次检索的准确率 A_c 计算^[13]如下:

$$A_c = \frac{N_1}{N} \times 100\% \quad (9)$$

其中, N_1 为检索准确的图像数目, N 为返回的总图像数目。

最后求出每次反馈后的平均检索准确率,作为评价本文方法检索效果的最终指标。

3.2 实验结果及分析

表 1 给出了每类图像初次检索及每次反馈后的平均检索准确率,以及各类图像的总平均检索准确率。图 3 为部分检索结果。程序初次检索耗时平均为 2.0s,反馈检索耗时平均为 1.9s。

表 1 各类型图像检索准确率对比

Tab 1 Retrieval accuracy comparison of each kind of images

	街区	植被	街区植被	道路植被	总平均
初次检索	81.40	100	38.40	57.00	69.20
1次反馈	84.60	100	59.20	67.00	77.70
2次反馈	89.00	100	73.60	84.00	86.65
3次反馈	92.00	100	82.00	89.00	90.75
4次反馈	90.80	100	84.00	87.80	90.65
5次反馈	92.00	100	83.20	88.20	90.85
6次反馈	92.40	100	83.60	87.40	90.85
7次反馈	91.60	100	83.20	87.80	90.75

由图 3 的直观显示及表 1 的统计数值分析可得以下结论:

1) 对于街区、植被等单一用途地类的初次检索平均准确率比混合用地较高。对植被的初次检索平

均准确率可以达到 100%，对街区的初次检索平均准确率可达到 81.14%。对街区植被混合用地的初次检索准确率仅为 38.40%，对道路植被混合用地的初次检索准确率仅为 57%。

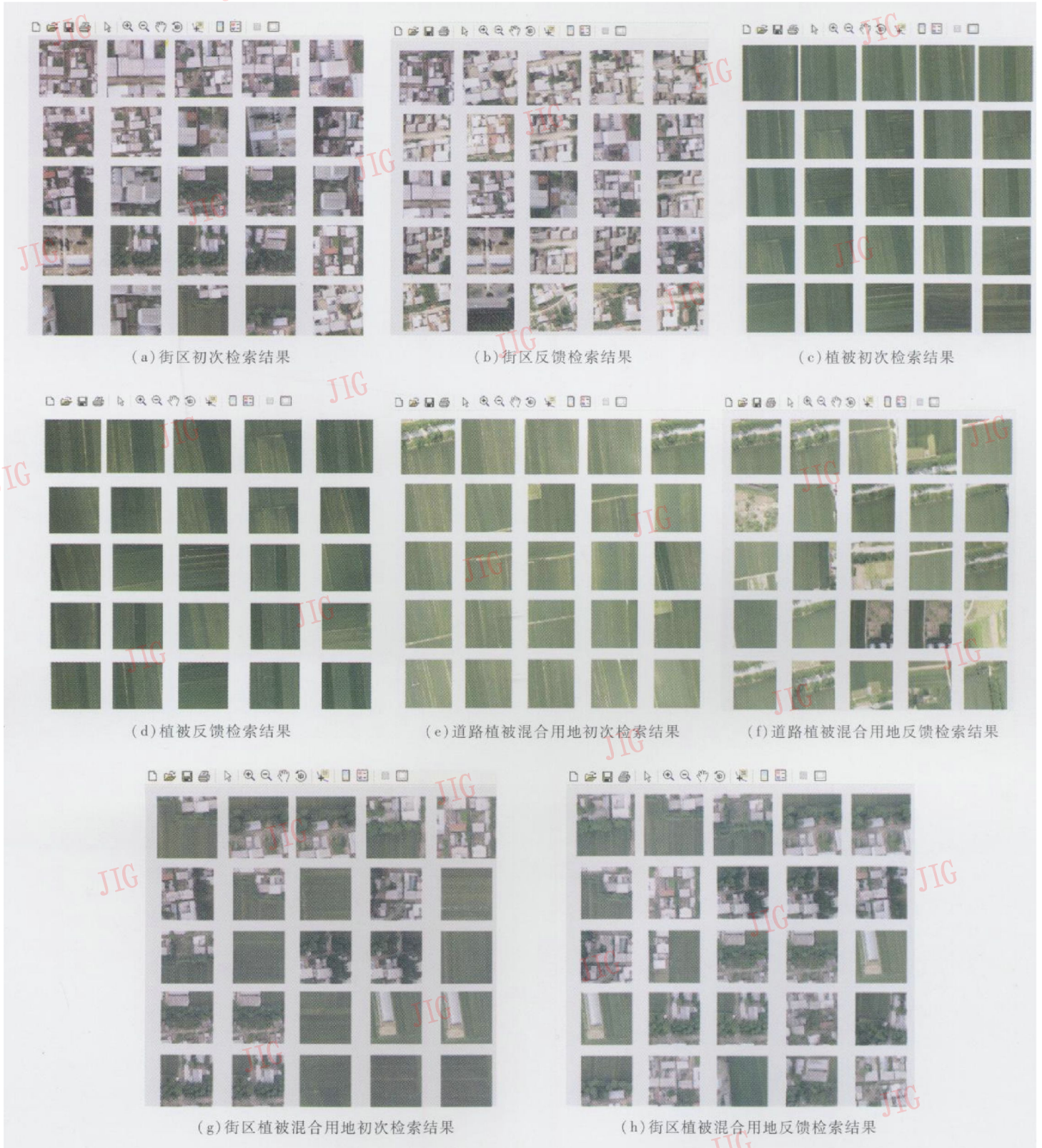


图 3 各类型图像初次检索及反馈检索部分实验结果

Fig. 3 Partial results of first time and feedback retrieval of each kind of images

2)经本文方法逐次反馈后,各地类影像的检索准确率逐渐上升,特别是混合用地的检索准确率提升显著,街区植被混合用地的初次检索平均准确率仅为 38.40%,经 3 次反馈后,平均准确率达到 82.00%,对道路植被混合用地的初次检索平均准确率为 57.00%,经 3 次反馈后,平均准确率达到 89.00%。

3)当反馈次数超过 3 次后,检索准确率趋于稳定,只有小幅波动。

实验结果证明,本文方法能够提高传统的基于小波纹理的图像检索算法的准确性。

4 结 论

本文将人工免疫原理应用于遥感图像检索,即根据图像检索的相关反馈技术及免疫机理,利用免疫算法的自动识别、无教师学习及记忆的特点,对用户检索的语义目标进行跟踪,使得检索结果越来越趋近于用户的语义要求,从而提高了传统的基于小波纹理检索算法的检索精度,但针对其他算法的适用性还有待进一步研究实验。由于人工免疫原理本身还处于研究和发展阶段,随着研究的不断深入,人工免疫算法的性能会不断提高,应用领域也会不断拓展,其在图像检索方面必将发挥更大的作用。

参考文献 (References)

- [1] Shen Yun-tao. Research on Image Retrieval Based on Human Perception Imitation [D]. Xi'an Northwest Polytechnical University, 2005 [沈云涛. 基于视觉感知特性的图像检索研究 [D]. 西安: 西北工业大学, 2005.]
- [2] Duan Fu, Zhang Ming-fan. Application of immune algorithm in image retrieval [J]. Control & Automation, 2007, 23(4-3): 279-280 [段富, 张明凡. 免疫算法在图像检索的应用 [J]. 微计算机信息, 2007, 23(4-3): 279-280.]
- [3] Jiao Li-cheng, Du Hai-feng, Liu Fang et al. Immune Computing Learning and Recognition Optimization [M]. Beijing: Science Press, 2006 [焦李成, 杜海峰, 刘芳等. 免疫优化计算、学习与识别 [M]. 北京: 科学出版社, 2006.]
- [4] Takana Y. WWW information visualization based on plastic clustering [C] // Jin Bezdek. Proceedings of IEEE International Fuzzy Systems Conference, Melbourne, Australia, IEEE Inc., 2001: 1295-1298.
- [5] Takana Y, Hirota K. Employing immune network model for clustering with plastic structure [C] // Zhang Hong. Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, Banff, Alberta, Canada, IEEE Inc., 2001: 178-183.
- [6] Takana Y, Hori T. Application of immune network metaphor to keyword map-based topic stream visualization [C] // Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, 2003: 770-775.
- [7] Zhao Meng-de. Text information retrieval model based on immunology construction [J]. Journal of East China University of Science and Technology (Nature Science), 2007, 33(1): 132-136 [赵孟德. 基于免疫学架构的文本信息检索模型 [J]. 华东理工大学学报 (自然科学版), 2007, 33(1): 132-136.]
- [8] Bumet F M. The Clonal Selection Theory of Acquired Immunity [M]. Cambridge, UK: Cambridge University Press, 1959.
- [9] De Castro L N, Von Zuben F J. Learning and optimization using the clonal selection principle [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(3): 239-251.
- [10] De Castro L N, Von Zuben F J. The clonal selection algorithm with Engineering Applications [C] // L. Darrel Whitley. Proceedings of Genetic and Evolutionary Computation Conference, San Francisco, CA, USA: Morgan Kaufmann Publishers, 2000: 36-37.
- [11] Wu Dong-sheng, Wu Le-nan, Huang Bo. Texture and non-texture images retrieval based on wavelet fuzzy clustering [J]. Journal of Image and Graphics, 2003, 8(12): 1400-1405 [吴冬升, 吴乐南, 黄波. 基于小波模糊聚类的均质纹理和非均质纹理图像检索 [J]. 中国图象图形学报, 2003, 8(12): 1400-1405.]
- [12] Lu Li-zhen, Liu Ren-yi, Liu Nan. Remote sensing image retrieval using color and texture fused features [J]. Journal of Image and Graphics, 2004, 9(3): 328-333 [陆丽珍, 刘仁义, 刘南. 一种融合颜色和纹理特征的遥感图像检索方法 [J]. 中国图象图形学报, 2004, 9(3): 328-333.]
- [13] Wang Wen-hui. Research on Content-based Image Retrieval Techniques [D]. Changsha: National University of Defense Technology, 2001: 32-62 [王文惠. 基于内容的图像检索技术研究 [D]. 长沙: 国防科学技术大学, 2001: 32-62.]