

Feature Weight Assignment Based on AIRS for Remote Sensing Image Retrieval

Lichuan Geng^{1,2*}, Zexun Geng^{1,2,3}

¹School of Urban-Rural Planning and Landscape Architecture, Xuchang University, Xuchang Henan

²Collaborative Innovation Center for UAVLRS, Xuchang University, Xuchang Henan

³The Information and Engineering University of the PLA, Zhengzhou Henan

Email: *glch1982@163.com

Received: Mar. 19th, 2015; accepted: Apr. 2nd, 2015; published: Apr. 8th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In view of the effective feature extraction issue in current image retrieval, a feature weight assignment method based on Artificial Immune Recognition System (AIRS) is promoted in this paper. The generalized learning and memory character of AIRS is used to learn the character of training samples, in order to determine the weight between the image features. Experimental results show that comparing to the traditional feature weight assignment method, this method can provide better feature weight, and improve the image retrieval accuracy.

Keywords

Remote Sensing Image Retrieval, Effective Feature, AIRS, Feature Weight

基于AIRS的遥感图像检索特征权值调整方法

耿利川^{1,2*}, 耿则勋^{1,2,3}

¹城乡规划与园林学院, 许昌学院, 河南 许昌

²无人机低空遥感技术协同创新中心, 许昌学院, 河南 许昌

³解放军信息工程大学, 河南 郑州

Email: *glch1982@163.com

*通讯作者。

收稿日期：2015年3月19日；录用日期：2015年4月2日；发布日期：2015年4月8日

摘要

针对当前图像检索中存在的有效特征提取问题，提出了一种基于人工免疫识别系统(AIRS)的特征权值调整方法。利用人工免疫识别系统的泛化学习及记忆的特点，对训练样本的特征值进行学习，从而确定各特征之间的权值分配。实验结果表明，与传统权值调整法相比，本方法能够为各特征提供较好的权值，提高图像检索的准确率。

关键词

遥感图像检索，有效特征，AIRS，特征权重

1. 引言

基于内容的图像检索技术作为当前遥感图像处理和计算机视觉研究的一个热点[1]-[3]，还面临着许多问题。其中一个突出问题就是如何提取图像的有效特征，自适应确定多维特征之间的权值分配。图像特征维数升高并不一定能够提高检索的准确性，相反有些特征的存在甚至影响检索的准确性。很多学者在这方面提出了一些解决方法。Rui 和 Huang [4]最早将相关反馈技术引入到图像检索中，提出了利用特征标准差的方法来调整权值，使相似性计算尽量与用户的感受一致，该方法虽然简单，并且具有一定的数学理论支持，但是需要用户反复交互，给系统的实用性方面带来了挑战[5]；汪祖媛等[6]提出了一种基于进化规划策略的纹理图像检索方法，利用进化规划具有群体搜索和随机信息交换的优点，对图像特征矢量进行优化选取，但是由于进化算法本身存在容易陷入局部极小点，收敛速度慢的缺点，使该方法不容易找到最优特征；崔文成等[7]提出了一种基于遗传算法的多特征权重自动赋值方法，首先将图像检索中特征权重调整问题转换为优化问题，进一步利用遗传算法求取权重的最优解，该算法同样存在算法不能保证概率收敛的问题[8]；王晓翠等[9]提出了一种基于改进的遗传算法用于解决商标图像检索中的多特征权值分配问题，针对标准的遗传算法收敛速度慢，容易产生“早熟”现象，将遗传算法中的轮盘赌选择与最优解保存法结合，并将染色体基因重新排列后进行交叉，提高了检索系统的效率和精度，但是该系统用遗传算法分配权之后并不具有自适应能力。

人工免疫系统(Artificial Immune System, 简称 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，提供噪声忍耐、无教师学习、自组织、记忆等进化学习机理，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、优化学习和故障诊断等许多领域[8]。文献[10]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，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，而且需要用户反馈的次数相对比较多次，算法精度收敛比较慢。

针对以上问题，本文借鉴了人工免疫识别系统(Artificial Immune Recognition System, AIRS)的基本思想[11]-[13]，将其引入到遥感图像检索的特征权值调整中，利用其自组织、记忆、泛化学习的特点，对特征权值进行调整，去除权值过小的特征，保留图像的有效特征，实现保留特征之间的权值最优分配。实验结果表明，本文方法是有效的。

2. AIRS 的基本理论与算法

2.1. AIRS 基本理论

生物免疫系统的主要任务是清除侵入体内的病原物质，并对病原体的再次入侵做好防御准备。免疫系统包括 B 细胞和 T 细胞两种淋巴细胞，每个淋巴细胞表面都存在受体分子，这些受体可以与抗原提呈细胞提呈的抗原分子按照一定的亲和力产生结合。受抗原刺激，具有抗原识别能力的淋巴细胞，通过克隆增殖、变异和进化，最后达到亲和力成熟，产生记忆细胞。AIRS 的训练便是借鉴免疫系统对于一个入侵抗原的响应过程。2001 年 Watkin 在 Timmis 提出的资源受限人工免疫系统的基础上，提出了人工免疫识别系统[11]，后又经几次改进[12] [13]。AIRS 的中心思想是：进入免疫系统的抗原，由与之最具亲和力的记忆抗体提呈。该记忆抗体经过克隆、变异产生出大量的新抗体，即人工免疫识别球(Artificial Recognition Ball, ARB)，ARB 中抗体之间再经过资源竞争，进化得到与抗原更具亲和力的抗体，更新记忆抗体集。最后得到一组能够很好识别抗原的记忆细胞。

2.2. AIRS 算法流程

AIRS 算法包括四个基本步骤：

1) 初始化

这个步骤可以看作一个数据预处理过程，主要包括特征向量的规则化以及记忆细胞集 MC 及 ARB 的生成。

下面是一个循环过程，对每个抗原 Ag_i 执行以下步骤：

2) 生成新的 ARB

找出与抗原 Ag_i 刺激度最大的记忆细胞 mc_{match} ，由它进行克隆增殖生成新的 ARB。

3) 资源竞争产生新记忆细胞

ARB 中抗体之间根据与抗原 Ag_i 亲和力的大小进行资源竞争，获胜的抗体称之为候选记忆细胞 mc_{cand} ，即有机会进入记忆细胞。

4) 候选记忆细胞进入记忆细胞集

如果候选记忆细胞 mc_{cand} 与抗原 Ag_i 的亲和力大于现有的记忆细胞，则将候选记忆细胞加入记忆细胞集 MC，同时如果候选记忆细胞 mc_{cand} 与 mc_{catch} 之间的亲和力大于用户设定的某个阈值，则 mc_{cand} 代替 mc_{catch} 进入记忆细胞集。

对所有抗原学习完毕后，循环结束。

3. 基于 AIRS 的特征权值调整方法

3.1. 特征权值调整方法基本过程

本文方法将同类样本特征向量之间的差值绝对值作为训练抗原，将特征权值作为抗体，特征权重调整的过程就是不断优化抗体，识别抗原，最后直到产生可以捕获抗原的最佳抗体过程。具体步骤如下：

1) 初始化

设训练样本中存在 M 幅样本图像，分为 C 类，记为 $\mathbf{X}_j^i (i=1,2,\dots,C; j=1,2,\dots,M)$ ，首先提取图像的特征向量 \mathbf{V}_j^i ，分别计算同类特征向量总体的均值 $\bar{\mathbf{V}}^i (i=1,2,\dots,C)$ ，然后计算特征向量与该类型特征均值差的绝对值，记为 $\mathbf{A}g_j (j=1,2,\dots,M)$ 作为训练抗原。将记忆抗体 $\mathbf{A}b_k (k=1,2,\dots,K)$ 初始化为一组与抗原长度相同，符合高斯分布的随机向量， K 为常量，抗体均值为 1，方差为 δ ，方差为一个小值。

2) 记忆抗体训练

由抗原对记忆抗体进行训练，寻找记忆抗体中亲和力最好的 N 个抗体，对其进行较高变异率的克隆变异，克隆数由其亲和力 $mc.aff$ 的大小及资源总数 $Totalres$ 决定，与最佳个体共同生成免疫识别球 ARB，重新计算抗原与 ARB 的亲和力，保留与之亲和力最高的 N 个 ARB 抗体，对其进行较低变异率的克隆变异，克隆数同样由其亲和力的大小及资源总数决定，计算重新生成 ARB 与抗原的亲和力，将亲和力最高的 H 个抗体代替记忆抗体中的低亲和力的 H 个抗体。如此循环直到对所有的抗原训练完毕。对记忆抗体集求平均值，将值小于阈值 ε 的值置零，作为最优记忆抗体。训练过程中主要包括以下几个运算：

3) 克隆运算

克隆运算即是简单的复制，克隆数由记忆抗体亲和力 $mc.aff$ 及资源总数 $Totalres$ 确定，算法如下：

$$mc_i.res = \frac{1/mc_i.aff}{\sum_i 1/mc_i.aff} \cdot Totalres \quad (1)$$

4) 变异运算

变异运算是一个改变抗体值的过程，文中将抗体对应相乘一个符合高斯分布的随机向量，向量均值为 1，方差为 ζ ，本文中有两次变异，初次为记忆细胞的变异， ζ_1 设置的值较大，避免算法陷入局部最优，二次变异为 ARB 变异， ζ_2 设置的值较小，避免算法不收敛。

5) 亲和力运算

文中抗体与抗原之间的亲和力由它们之间的内积确定。

$$mc_i.aff = mc_i \bullet ag_j \quad (2)$$

6) 图像检索

利用生成的最优抗体，在计算目标图像与图像库中图像之间相似度时，进行加权计算，距离采用加权欧式距离。

3.2. 特征提取

3.2.1. 纹理特征提取

本文提取了图像的小波纹理特征，利用小波变换对图像进行三级小波分解，在子带 B (LL, LH, HL, HH) 中的第 mn (m 为行号， n 为列号) 个小波系数，在分解级数为 k 时记为 $(c_{m,n}^k, d_{m,n}^{k1}, d_{m,n}^{k2}, d_{m,n}^{k3})$ ，系数的能量记为[14]： $e_{m,n}^k = c_{m,n}^k \cdot c_{m,n}^k$ 。提取各子带能量的一阶矩、二阶矩为特征。数学上前二阶矩定义如下：

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

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

提取特征后，每幅彩色图像生成 60 维特征向量，记为： $V_i (i = 0, 1, \dots, M - 1)$ ， M 为图像数。

3.2.2. 颜色特征提取

图像的颜色有多种表示方式，其中 HSV 模型对应于画家配色模型，能较好反应人对颜色的感知和鉴别能力是一种适合人们肉眼分辨的模型。此外，它还具有线性伸缩性，可感知的颜色与颜色分量的相应值得欧氏距离成比例，因此非常适合于颜色的图像相似性比较[15]。

在本文中，我们将 HSI 色彩空间非均匀量化成 72 维，即把色调 Hue 分成八份(人眼对 Hue 最敏感)，饱和度 S 和亮度 I 各分成三份。具体量化值如下式：

$$H = \begin{cases} 0, h \in [316, 20] \\ 1, h \in [21, 40] \\ 2, h \in [41, 75] \\ 3, h \in [76, 155] \\ 4, h \in [156, 190] \\ 5, h \in [191, 270] \\ 6, h \in [271, 295] \\ 7, h \in [296, 315] \end{cases} \quad S = \begin{cases} 0, s \in [0, 0.2] \\ 1, s \in [0.2, 0.7] \\ 2, s \in [0.7, 1] \end{cases}$$

$$V = \begin{cases} 0, s \in [0, 0.2] \\ 1, s \in [0.2, 0.7] \\ 2, s \in [0.7, 1] \end{cases}$$

量化后，根据(5)式将图像转换为灰度图像，图像灰度 G 取值范围为 $[0, 1, \dots, 71]$ ，计算该灰度图像的累加直方图，作为图像的颜色特征。

$$G = 9H + 3S + I \quad (5)$$

4. 实验结果与分析

本文使用的实验图像库由 2005 年 5 月 8 日，用无人机搭载 Canon 1DsMark-II 相机，在驻马店拍摄的影像，从中抽取 1000 幅 512×512 的子图组成，图像内容包括街区，道路和植被三类区域。本文进行了两组实验，第一组只利用小波纹理进行检索，第二组组合小波纹理和颜色纹理进行检索，并对利用本文方法进行权值优化前后结果与文献[4]提出的利用特征标准差调整权值方法的实验结果进行了对比。

4.1. 小波纹理检索

利用 db4 小波对训练图像进行三层小波变换，提取系数 60 维特征向量，初始化记忆细胞集以及抗原，记忆细胞集为 60 维均值为 1，方差为 0.01 的向量组，个数为 30。抗原为各特征向量与同类特征向量均值差的绝对值。

本算法在计算机处理器为双核主频 2.10 GHz、内存 2 G、Windows Vista、matlab 7.0 环境中实现。实验中参数设置如下：资源总数为 1000，记忆变异时， ζ_1 设为 0.1，ARB 变异时， ζ_2 设为 0.01，权值阈值 ε 设为 0.4，小于 ε 的权设为 0。

分别从每类图像中选取 10 幅有代表性的图像进行检索，显示前 25 个结果输出。每次检索的准确率由下式计算：

$$C = \frac{\text{返回的正确图像的数目}}{\text{返回的图像数目}} \times 100\% \quad (6)$$

最后求出不加权和本文方法的平均准确率，并与利用特征标准差调整权值的结果进行对比，作为评价本文方法的最终指标。

表 1 给出了三种方法检索的平均准确率及总平均准确率。图 1 为部分检索结果。图 1 中显示图像次序按照与待检索图像间亲和力的大小，由大到小依次排列。从显示检索结果对比可以看出，针对不同类型的土地利用类型，本文算法检索精度相比对图像特征均赋相同权值及 Rui 法检索结果均有一定程度的提高。

4.2. 组合小波纹理与颜色累加直方图的检索

将小波纹理特征和颜色累加直方图特征组合成图像的混合特征进行检索，对特征权值优化前后检索结果进行对比。参数设置与 4.1 中参数设置相同。表 2 给出了两种方法检索的平均准确率及总平均准确率。图 2 为部分检索结果。图 2 中图像排列顺序与图 1 相同，检索结果对比与图 1 中的结果对比一致。

Table 1. Wavelet feature image retrieval accuracy comparison
表 1. 小波特征图像检索准确率对比

| | 植被 | 街区 | 街区植被 | 植被道路 | 总平均准确率 |
|---------|-------|-------|-------|-------|--------|
| 不加权结果 | 81.2% | 74.2% | 68.0% | 66.5% | 72.5% |
| Rui 法结果 | 83.2% | 76.4% | 71.6% | 72.4% | 75.9% |
| AIRSIR | 89.6% | 88.1% | 79.6% | 73.0% | 82.7% |

Table 2. Wavelet and color combination feature image retrieval accuracy comparison
表 2. 小波和颜色组合特征图像检索准确率对比

| | 植被 | 街区 | 街区植被 | 植被道路 | 总平均准确率 |
|---------|--------|-------|-------|-------|--------|
| 不加权结果 | 79.1% | 80.9% | 78.2% | 84.4% | 80.7% |
| Rui 法结果 | 84.5% | 82.0% | 81.6% | 85.6% | 83.5% |
| AIRSIR | 93.33% | 86.2% | 86.7% | 93.3% | 89.9% |

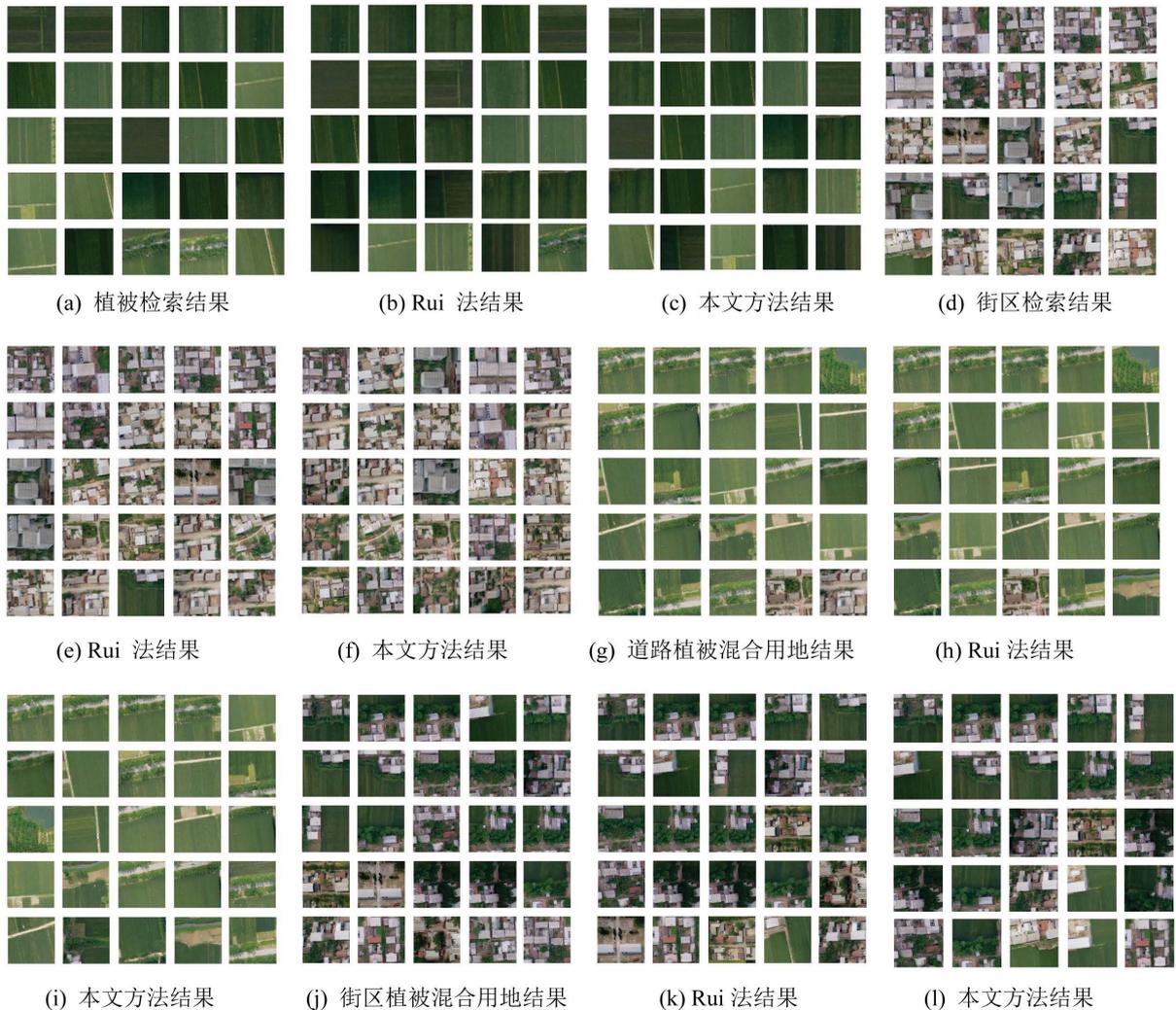


Figure 1. Wavelet feature image retrieval partial experiment results
图 1. 小波特征图像检索部分实验结果

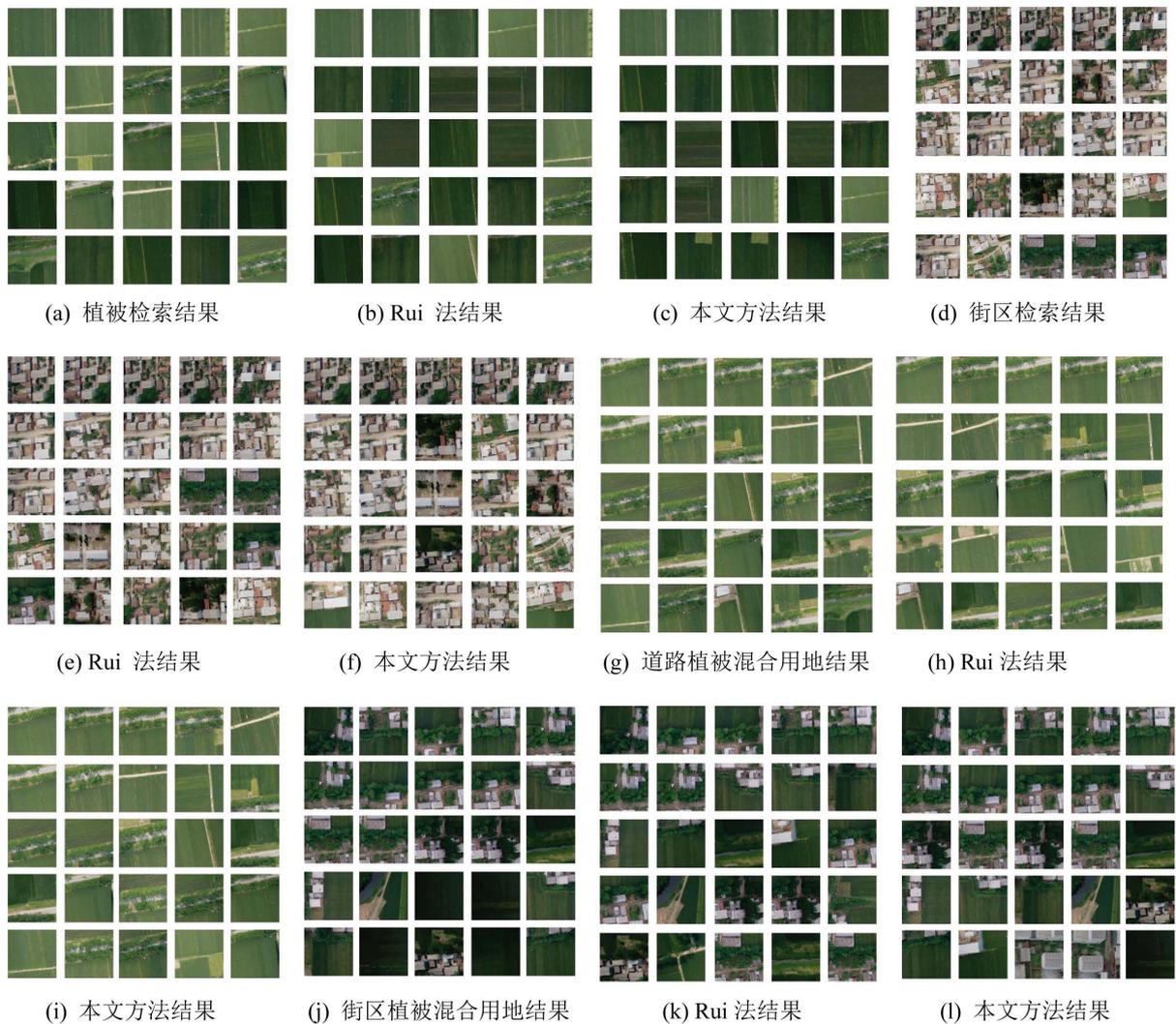


Figure 2. Wavelet and color combination feature image retrieval partial experiment results

图 2. 小波和颜色组合特征图像检索部分实验结果

4.3. 实验分析

实验对植被、居民地、道路植被混合用地、街区植被混合用地四类遥感图像进行了检索。

由图 1 和图 2 以及统计表对比可以得出以下结论：

1) 利用特征标准差调整权值的方法和本文方法对特征权值进行优化调整之后，图像检索的程度整体都有所提高，但是本文方法的检索准确率要优于 Rui 法；

2) 单一小波纹理的检索精度要低于小波纹理与颜色累加直方图组合特征的检索精度，经过特征权值优化之后，整体检索结果不变，只是植被组合特征的检索结果要稍微低于单一特征的检索结果，这与一些偶然因素有关，如人工判断因素的影响；

3) 由表 1 及图 1 可以看出，单一小波纹理的检索，对于单一土地类型的检索精度要高于混合土地类型的检索精度。经多特征权值优化之后，结果不变，原因与小波特征对图像内容表达能力有关；

4) 由表 2 及图 2 可以看出，混合特征的检索对土地类型是单一还是混合不敏感，所以在提取特征的时候，最好提取图像的多个特征进行组合。

5. 结论

本文将 AIRS 应用于遥感图像检索的多特征权值分配优化问题中, 利用免疫算法的泛化学习、记忆及收敛快速的特点, 对特征权值进行优化, 很好的解决了特征权值自适应调整问题, 提高了图像检索精度。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, 在图像检索方面必将发挥更大的作用。

参考文献 (References)

- [1] 耿利川 (2008) 基于人工免疫系统的遥感图像检索算法研究. 解放军信息工程大学, 郑州.
- [2] 李德仁, 宁晓刚 (2006) 一种新的基于内容遥感图像检索的图像分块策略. 武汉大学学报·信息科学版, **31**, 659-662.
- [3] Bao, Q. and Guo, P. (2004) Comparative studies on similarity measures for remote sensing image retrieval. *IEEE International Conference on Systems, Man and Cybernetics*, Hague.
- [4] Rui, Y., Huang, T.S. and Mehrotra, S. (1998) Relevance feedback: A powerful tool in interactive content-based image retrieval. *IEEE Transactions on CSVT*, **8**, 644-655.
- [5] 韦娜 (2006) 基于内容的图像检索技术. 西北大学, 13-14.
- [6] 汪祖媛, 李斌, 罗琳, 庄镇泉 (2001) 基于进化规划策略的纹理图像检索. *小型微型计算机系统*, **22**, 950-953.
- [7] 崔文成, 邵红 (2008) 基于遗传算法的图像检索中特征权重自动调整. *计算机工程与应用*, **44**, 106-108.
- [8] 焦李成, 杜海峰, 刘芳, 公茂果 (2006) 免疫优化计算、学习与识别. 科学出版社, 51.
- [9] 王晓翠, 刘建平 (2008) 基于改进遗传算法的商标多特征权值分配. *计算机工程与设计*, **29**, 2038-2040.
- [10] 段富, 张明凡 (2007) 免疫算法在图像检索的应用. *微计算机信息*, **23**, 279-280.
- [11] Watkins, A. (2002) AIRS: A resource limited artificial immune classifier. Master's Thesis, Mississippi State University.
- [12] Watkins, A. and Timmis, J. (2002) Artificial Immune Recognition System (AIRS): Revisions and refinement. In: Timmis, J. and Bentley, P.J., Eds., *Proceedings of the International Conference on Artificial Immune Systems (ICARIS)*, 99-106.
- [13] Watkins, A. and Boggess, L. (2004) Artificial Immune Recognition System (AIRS): An Immune-Inspired Supervised Learning Algorithm. *Genetic Programming and Evolvable Machines*, **5**, 291-317.
- [14] 吴冬升, 吴乐南, 黄波 (2003) 基于小波模糊聚类的均质纹理和非均质纹理图像检索. *中国图象图形学报*, **8A**, 1400-1405.
- [15] Smith, J.R. and Chang, S.-F. (1996) Tools and Techniques for Color Image Retrieval. *SPIE*, **2670**, 426-437.