

一种基于生物群集智能优化的遥感分类方法

刘小平^① 黎夏^{①*} 彭晓鹃^② 黎海波^① 何晋强^①

(① 中山大学地理科学与规划学院, 广州 510275; ② 南海环境监测中心, 广州 510300)

摘要 提出了一种基于生物群集智能(swarm intelligence-based)的遥感影像分类新方法. 智能式分类是遥感研究的热点和趋势. 设计出一种新的实数制编码的粒子群遥感影像分类器(PSO-Miner), 在分类规则提取时, 粒子能自动寻找各个波段的最优分割点. 并且该方法所提取的分类规则毋需通过数学公式来表达, 能更方便和准确地描述自然界中的复杂关系, 比数学公式更容易让人理解. 将该方法应用于番禺地区的遥感影像, 取得了较好的分类结果. 并与 See5.0 决策树方法进行了对比研究, 实验结果表明, 基于群集智能方法的分类精度比决策树方法的精度更高.

关键词 群集智能 粒子群(Particle Swarm Optimization, PSO) 遥感影像 分类

影像分类是将遥感数据转化为便于使用的地理数据的核心^[1]. 传统的分类方法包括最短距离、最大似然、聚类分析、Bayesian分类等^[2,3], 这些分类方法简单实用, 但都是基于统计学原理的, 以训练数据遵循正态分布为前提, 因此训练样本的选择和参数估计直接影响分类结果^[4]. 最近, 学者们提出了许多新的遥感分类方法, 包括机器学习^[5,6]、支持向量机(SVM)^[7,8]以及基于计算智能(CI)的神经网络^[9,10]、模糊集^[11]、遗传算法^[12]等. 这些方法在一定程度上提高了分类的精度, 但是, 遥感分类精度还有较大的改善空间^[13]. 因此, 如何提高土地利用分类精度仍是遥感领域一个重要研究方向.

目前, 随着人工智能技术的发展, 遥感分类也趋向与智能化技术相结合. 生物群集智能作为一种自下而上的研究方法, 其本质是一个复杂的多智能体系统, 它由众多简单个体(如蚂蚁、鸟类)所组成, 通过个体间相互协作和竞争而表现出群集智能行为.

生物群集智能目前已经成为人工智能研究的热点, 并在旅行商问题求解、数据聚类、组合优化、网络路由、模式识别等方面取得了系列较好的成果^[14~18]. 可惜, 群集智能在遥感分类方面的研究非常有限. 本文尝试将生物群集智能中的粒子群算法引入到遥感图像分类中, 并进行了探索性的研究. 粒子群算法是由 Kennedy 和 Eberhart^[19]提出, 源于对鸟群觅食行为的研究. 其最初的设想是仿真简单的社会系统, 并解释复杂的社会行为, 却发现 PSO 是解决复杂优化问题的有效技术^[20]. 与进化算法相比较, PSO 保留了基于种群的全局搜索策略, 采用简单的速度-位移模型, 避免了复杂的遗传操作, 同时它的记忆能力使其可以跟踪当前的搜索情况并动态调整搜索策略, 具有较强的全局收敛能力和鲁棒性^[21]. 因此, PSO 作为一种高效的并行搜索算法, 非常适于地理复杂问题的求解.

当地表状况较复杂时, 目前的遥感分类方法并

收稿日期: 2007-05-21; 接受日期: 2007-07-18

国家杰出青年基金项目(批准号: 40525002)、国家高技术研究发展计划(编号: 2006AA12Z2062)、国家海洋局青年海洋科学基金(批准号: 2007430)和南海分局海洋科学技术局长基金(批准号: 0641)资助

* 联系人, E-mail: lixia@mail.sysu.edu.cn

不能很好地处理线性不可分的地物, 智能式的分类方法或许能够为解决这一问题提供新的手段. 本文以土地利用遥感分类为例, 提出了利用粒子群分类器(PSO-Miner)对遥感影像进行分类的新方法. 其原理是通过模仿鸟群觅食行为的方式来构造遥感分类的规则. 在 PSO-Miner 算法中, 粒子在多维空间中, 不断地搜索各波段的最优区间, 以构造分类规则, 各个波段的最优区间与土地利用类别连接, 就形成一条规则. 通过序列覆盖算法移除规则所覆盖的训练数据, 如果训练数据中某一类别的数据个数小于一个阈值时, 可终止该类别数据的规则挖掘, 再利用 PSO 算法进行下一类别规则的挖掘, 直至所有类别的规则挖掘完毕. 基于 PSO 的遥感分类方法所提取的 If-Then 分类规则并不需要通过数学公式来表达, 能更方便和准确地描述自然界中的复杂关系.

1 粒子群算法的基本原理

在 PSO 中, 每个优化问题的潜在解相当于搜索空间中一只鸟, 称之为粒子, 所有的粒子都有一个被优化函数决定的适应值, 每个粒子的速度向量决定它们飞翔的方向和距离. PSO 初始化为一群随机粒子, 通过迭代找到最优解. 粒子通过跟踪两个极值来更新自己, 第一个极值就是粒子本身到当前时刻为止找到的最好解, 这个解称之为个体极值($pbest$); 另一个极值就是整个种群到当前时刻找到的最好解, 这个值称为全局极值($gbest$). 粒子追随当前的最优粒子在解的空间中进行搜索, 直至找到最优解^[19]. 每个粒子具有记忆的功能, 它可以记录前一次的结果, 粒子之间的合作和竞争能够表现出复杂的特性.

图 1 以二维空间为例, 展示了粒子在解的空间中进行优化搜索时的情况. 在图 1 中, X_t 表示粒子上一代的搜索位置, X_{t+1} 为粒子移动后的搜索位置; 粒子的移动速度为 3 个速度的矢量和, 它们分别是: V_t —上一次迭代的移动速度, V_{pbest} —粒子个体极值的移动速度, V_{gbest} —粒子全局极值的移动速度. 在速度 V_t , V_{pbest} , V_{gbest} 的共同作用下, 粒子将以速度 V_{t+1} 到达新的位置 X_{t+1} , 向全局最优解靠近. 如此迭代下去, 粒子将逐步逼近最优解, 并最终找到最优解或近似最优解.

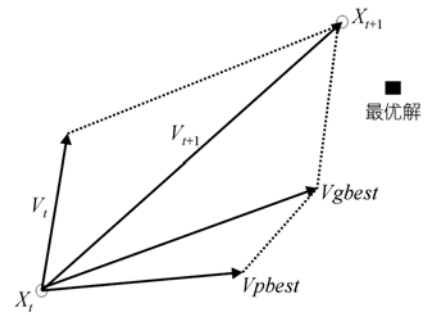


图 1 粒子优化搜索示意图

2 基于粒子群的遥感分类方法

粒子群算法具有较好的记忆能力, 可以跟踪当前的搜索情况并动态调整搜索策略. 粒子之间的协作和竞争关系使其具有较强的全局收敛能力和鲁棒性. 因此, 粒子群算法可以有效地解决非线性问题, 特别适合于地理复杂现象. 本文将以太空影像为例, 探讨利用粒子群分类器(PSO-Miner)对遥感影像进行分类.

PSO 在 TSP 问题求解、数据聚类、组合优化、模式识别等方面得到了广泛的应用^[14-16,18], 但在分类规则挖掘方面的研究非常有限. Sousa 等^[22]在 2003 年提出了基于 PSO 的规则挖掘算法, 但该算法是以二进制为编码的^[23]. 本文将采用一种实数制编码的 PSO 遥感影像分类器(PSO-Miner), 粒子能自动寻找遥感数据各波段的最优分割点. 二进制编码方式比较适合基于交叉和变异的算法, 如遗传算法, 但并不是特别适于粒子群算法. 在粒子群算法中, 每个粒子的位置和速度为一定范围的实数, 并且在优化过程中, 需要对粒子的速度和位置进行更新操作, 本文采用的实数制编码方式对粒子进行更新操作更加方便、简单和实用. 在 PSO-Miner 算法中, 每个粒子对应一条路径(图 2), 相应产生一条分类规则, 分类规则的挖掘可以当作是在多维空间对最优解的搜索. 规则的形式如下:

$$\begin{aligned}
 & \text{IF} \quad \text{band } 1 = \text{Value_1} \\
 & \quad \text{AND} \quad \text{band } 2 = \text{Value_2} \\
 & \quad \quad \quad \vdots \\
 & \quad \text{AND} \quad \text{band } j = \text{Value_}j \\
 & \text{THEN} \quad \text{Class_}x
 \end{aligned}$$

PSO 将每个解看成是在一个 D 维搜索空间内没有体积的粒子. 在本文中, 对于每一条规则, 粒子需

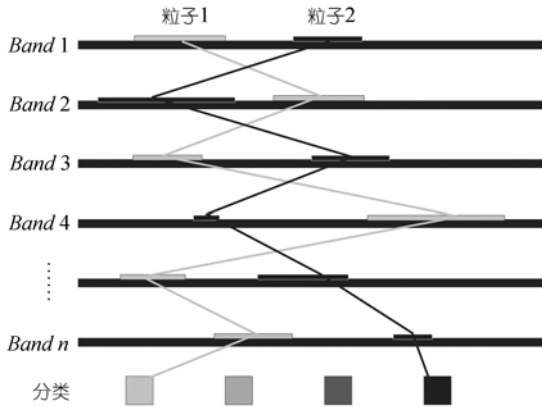


图 2 基于粒子群的遥感影像分类原理

要确定各波段的最优区间范围 $[x_-, x_+]$, x_- 为最优区间的下界, x_+ 为最优区间的上界(图 2), 即粒子自动寻找各个波段的最优分割点. 这种自动寻优与训练数据的类别数有一定关系, 训练数据类别数越多, 或者训练数据越复杂, 都会导致波段的分割点增多, 规则越复杂. 由于区间的上下限是成对出现的, 因此, 对于波段数目为 n 的遥感数据, $D=2n$. 假设有 m 个粒子, 第 i 个粒子的位置向量可表示为: $(x_{-i1}, x_{+i1}, x_{-i2}, x_{+i2}, \dots, x_{-in}, x_{+in})$, 第 i 个粒子的速度向量可表示为: $(v_{-i1}, v_{+i1}, v_{-i2}, v_{+i2}, \dots, v_{-in}, v_{+in})$. 粒子在飞行过程中不断根据当前最优值 $p(t)$ 和全局最优局 p_g 进行速度和位置的调整. 其速度调整函数如下^[19]:

$$\begin{cases} v_{-ij}(t+1) = w(t)v_{-ij}(t) + c_1 r_{1ij}(t)(p_{-ij}(t) - x_{-ij}(t)) + c_2 r_{2ij}(t)(p_{g-} - x_{-ij}(t)) \\ v_{+ij}(t+1) = w(t)v_{+ij}(t) + c_1 r_{3ij}(t)(p_{+ij}(t) - x_{+ij}(t)) + c_2 r_{4ij}(t)(p_{g+} - x_{+ij}(t)) \end{cases} \quad (1)$$

位置调整函数为:

$$\begin{cases} x_{-ij}(t+1) = x_{-ij}(t) + v_{-ij}(t+1) \\ x_{+ij}(t+1) = x_{+ij}(t) + v_{+ij}(t+1) \end{cases} \quad (2)$$

其中, $i=1, 2, \dots, m$; $j=1, 2, \dots, n$; t 为迭代次数. $w(t)$ 为第 t 次迭代的惯性权重, 用于控制前一次迭代产生的粒子速度对本次迭代速度的影响; $w(t)$ 较大时, 前一次迭代的粒子速度影响较大, 全局搜索能力较强, $w(t)$ 较小时, 前一次迭代的粒子速度影响较小, 局部搜索能力较强. c_1, c_2 是学习因子, 为非负常数, c_1 代表粒子的认知功能, 当 $c_1=0$ 时, 粒子没有认知功能, 容易陷入局部最优值; c_2 代表粒子之间的社会信息共享能力, 当 $c_2=0$ 时, 粒子之间没有信息共享; $r_{1ij}, r_{2ij}, r_{3ij}$,

r_{4ij} 为介于 $[0,1]$ 之间的随机数; $p_{-ij}(t), p_{+ij}(t)$ 为第 i 个粒子的迄今为止搜索到的遥感数据各波段下界和上界的适应度最优值; p_{g-}, p_{g+} 为整个粒子群迄今为止搜索到的遥感数据各波段下界和上界的适应度最优值, 其中, $p_{g-} = (p_{g-1}, p_{g-2}, \dots, p_{g-n})$, $p_{g+} = (p_{g+1}, p_{g+2}, \dots, p_{g+n})$. x_{-ij} 和 x_{+ij} 分别为粒子的下界和上界, 即粒子的位置. 当粒子搜索到最优解时, x_{-ij} 和 x_{+ij} 即为粒子自动寻找到的各个波段的最优分割点.

基于粒子群的遥感分类方法可分为 3 个阶段: 规则构造、分类规则适应度的计算、规则对训练数据的覆盖.

2.1 规则构造

规则构造模仿了鸟类的觅食行为. 粒子在每一个波段搜索最优上界和最优下界, 对于波段数目为 n 的遥感数据, 则是在 $2n$ 维空间内搜索最优值. 每个波段的最优区间用操作符‘And’连接, 与土地利用分类的类别相连形成一条规则.

粒子初始时随机分布在 $2n$ 维空间内, 初始状态粒子的位置分布为:

$$\begin{cases} x_{-ij} = Rand * (band_{jmax} - band_{jmin}) + band_{jmin} \\ x_{+ij} = Rand * (band_{jmax} - band_{jmin}) + band_{jmin} \end{cases} \quad (3)$$

其中, x_{-ij}, x_{+ij} 为第 i 个粒子的在第 j 个波段上的下界和上界, 如果 $x_{-ij} > x_{+ij}$, 则相互置换. $Rand$ 为分布在区间 $[0,1]$ 内的随机数. $band_{jmin}, band_{jmax}$ 分别为第 j 个波段的最大值和最小值. 粒子的初始速度为:

$$\begin{cases} v_{-ij} = Rand * v_{-j}^{max} \\ v_{+ij} = Rand * v_{+j}^{max} \end{cases} \quad (4)$$

v_{-j}^{max} 和 v_{+j}^{max} 分别为下界和上界方向上的最大速度.

运行后, 计算每个粒子的适应度值, 比较每个粒子的当前适应度(fitness)值与迭代前的个体最优值, 若粒子的当前适应度优于迭代前的个体最优值, 则进行个体最优值的更新. 否则个体最优值将不进行更新; 计算所有粒子的个体最优值, 适应度最优的个体最优值即为全局最优值. 随后, 对公式(1)中的惯性权重进行更新, 惯性权重更新的公式如下:

$$w(t) = w_{max} - t \cdot (w_{max} - w_{min}) / I_{max} \quad (5)$$

其中, t 为迭代次数, w_{\max} 为惯性权重最大值, w_{\min} 为惯性权重最小值, I_{\max} 为最大迭代次数.

根据粒子当前的个体最优值和全局最优值以及惯性权重, 利用公式(1)和(2)更新每个粒子的飞行速度和位置. 当全局最优适应度值与平均适应度值的绝对值小于一个阈值或迭代次数超过最大迭代次数 I_{\max} 时, 循环终止, 得到一组分类规则. 规则构造的伪代码如下表 1.

2.2 分类规则的适应度

分类规则(粒子)的适应度(fitness)用来衡量粒子位置的优劣, 是判断粒子飞行方向的指南针, 合理地选择适应度函数对问题的求解有着重要的作用, 本文用下面公式计算分类规则(粒子)的适应度:

$$Q = \left(\frac{\text{TruePos}}{\text{TruePos} + \text{FalseNeg}} \right) \cdot \left(\frac{\text{TrueNeg}}{\text{FalsePos} + \text{TrueNeg}} \right), \quad (6)$$

式中, TruePos 表示满足规则条件, 并且和规则预测类型相同的样例数; FalsePos 表示满足规则条件, 并且和规则预测类型不同的样例数; FalseNeg 表示不满足规则条件, 并且和规则预测类型相同的样例数; TrueNeg 表示不满足规则条件, 并且和规则预测类型不同的样例数.

2.3 规则对训练数据的覆盖

将搜索到的粒子最优位置 p_g (最优分类规则)置入到规则集 R 中, 然后采用序列覆盖算法在训练数据中移除规则 p_g 所覆盖的数据, 即波段属性和类别属性均与规则相匹配的数据, 其他类别属性的数据得以保留. 如果训练数据中某一类别的数据个数小于一个阈值时, 意味着该类别的数据量过小, 可终止该类别数据的规则挖掘, 进行下一类别规则的挖掘, 直至所有类别的规则挖掘完毕.

3 影像分类实验

实验中采用的数据是 2004 年获取的广州市番禺地区的 TM 卫星数据, 选取的研究区域范围为 1666×2211 个像元(分辨率为 30 m), 选择的波段为 TM 的 1~5 波段及 7 波段, 总共 6 个波段的数据. 图 3 为研究区域 5,4,3 波段所合成的假彩色影像图.

训练样本的选择是遥感影像分类的关键步骤, 直接关系到所获规则的质量. 根据遥感影像和实地调查来采集训练样本和验证数据集. 训练数据集的样本数为 2120, 验证数据集的样本数为 2000.

基于粒子群的遥感分类模型主要由两部分组成: 分类规则提取及遥感影像土地利用类型的识别. 分

表 1 基于粒子群的分类规则构造的伪代码

<pre> Input training data Initialize particle swarm /*初始化粒子群*/ While (t<Max_iteration or err>Min_error_criteria) /* 当 t 小于最大迭代次数或误差大于最小误差标准时*/ For i = 1 to No_of_particles /*粒子数目*/ Calculate particle's down_fitness /*计算粒子下界的适应度*/ If the down_fitness value is better than the best down_fitness value (p_best) in history Then Set current value as the new p_best End if Choose the particle with the best down_fitness value of all the particles as the g_best Calculate particle's upper_fitness /*计算粒子上界的适应度*/ If the upper_fitness value is better than the best upper_fitness value (p_+best) in history Then Set current value as the new p_+best End if Choose the particle with the best upper_fitness value of all the particles as the g_+best Update particle velocity according to equation (1) /*更新粒子的速度*/ Update particle position according to equation (2) /*更新粒子的位置*/ Next i Loop </pre>
--



图 3 番禺 TM 影像合成图

类规则主要通过 PSO-Miner 算法从训练数据挖掘出来的, PSO-Miner 算法是在 Visual Basic 6.0 环境中编程所实现的. 土地利用类型的识别部分则是根据 PSO-Miner 算法所挖掘出来的分类规则对遥感影像进行分类, 也是通过在 Visual Basic 6.0 环境中编程实现的.

利用 PSO 进行分类规则的提取时, 首先需要设定各参数值. 在本文的遥感分类实验中, 各参数值的设置见图 4. 选择 TM 数据的 1~5 及 7 波段等 6 个波段参与分类, 由于需要确定波段的最优上界和最优下界, 因此, 粒子相当于在 $D = 2 * 6$ 维的空间内进行优化搜索. 遥感分类的类别数为 8 类. 每一类别的种群规模 (粒子数目, $Numb$) 设为 20, 粒子最大速度 (v_{max}) 为 10; 根据 Shi 等^[24]的研究结果, 较大的惯性权重有利于跳出局部极值, 而较小的惯性权重有利于算法的收敛, 因此, 可将惯性权重设为随时间线性减小. 惯性权重最大值 (w_{max}) 通常取 0.9, 惯性权重最小值 (w_{min}) 通常取 0.4, 本实验的惯性权重也是如此设置; 当迭代次数超过 100 次时, 将强行性终止循环. 当每一类剩余的样例数目 < 5 时, 则终止该类别数据的规则挖掘, 进行下一类别规则的挖掘; PSO 的学习因子的权重对算法的性能影响较大, c_1 代表粒子的认知功能, 当

$c_1=0$ 时, 粒子没有认知功能, 在粒子的相互作用下, 有能力到达新的搜索空间, 但容易陷入局部最优值; c_2 代表粒子之间的社会信息共享能力, 当 $c_2=0$ 时, 粒子之间没有信息共享, 得到最优解的概率很低. 本实验中, 我们取 $c_1=c_2=2$.



图 4 PSO 遥感影像分类参数设置

根据前面所选取的训练数据, 本实验利用 PSO-Miner 算法共获得了 40 条分类规则, 表 2 列出了部分的分类规则. 根据所获得的分类规则, 对实验区遥感影像进行分类, 分类结果见图 5(a). 同时, 我们选取相同的训练数据, 用 See5.0 决策树方法对实验区的遥感影像进行了分类, 分类结果见图 5(b).

为了能更清楚地进行两者分类结果的对比, 分别对两种方法分类的结果图做了局部放大处理, 见图 6. 对比图 6(b)、6(c)、6(e)、6(f) 的分类结果, 容易发现: See5.0 决策树方法将图 6(c)、6(f) 中的部分基塘错分为水体, 而 PSO 对基塘的识别则基本正确; 在图 6(f) 中, See5.0 决策树方法将部分休耕地错分为建成区. 此外, PSO 分类后的斑块比较均一, 更接近土地利用制图的要求. 可见, PSO 遥感分类结果较真实地反映了实际的土地利用覆盖类型.

为了对比两种方法的精度, 用先前选取的验证数据对粒子群 (PSO) 分类方法和 See5.0 决策树方法进行了精度测试, 将精度评价结果分别表示为混淆矩阵, 见表 3 和 4. 通过比较混淆矩阵, 粒子群 (PSO) 分类方法的总体精度为 84.6%, See5.0 决策树方法的总体精度为 81.8%, 总精度在评价由于机遇造成的一致性和实际观测的一致性之间的差别时会有所偏差^[25]. Kappa 系数则能有效地避免这种偏差, 尤其对于地理数据来说, Kappa 系数能从位置、数量和综合信息来评价模型的精度. 因此, 用 Kappa 系数来度量遥感分类的精度. Kappa 系数的计算公式如下.

表 2 PSO-Miner 所挖掘的部分分类规则

<p>Rule 1: IF $96.7 < B_1 < 141.7 \ \& \ 48.9 < B_2 < 73.9 \ \& \ 64.5 < B_3 < 118.4 \ \& \ 81.4 < B_4 < 103.3 \ \& \ 110.8 < B_5 < 150.3 \ \& \ 46.9 < B_7 < 98.4$ Then class=Urban</p> <p>Rule 2: IF $84 < B_1 < 89.7 \ \& \ 37 < B_2 < 48.2 \ \& \ 42 < B_3 < 62.9 \ \& \ 20 < B_4 < 41.2 \ \& \ 12.8 < B_5 < 45.3 \ \& \ 4.4 < B_7 < 23.9$ Then class=Water</p> <p>Rule 3: IF $77.4 < B_1 < 90.8 \ \& \ 34.9 < B_2 < 44.8 \ \& \ 36.5 < B_3 < 56.7 \ \& \ 81.4 < B_4 < 140 \ \& \ 59.3 < B_5 < 97 \ \& \ 18.4 < B_7 < 46.5$ Then class=Agriculture</p> <p>Rule 4: IF $107.5 < B_1 < 152.6 \ \& \ 58.9 < B_2 < 105.6 \ \& \ 89.7 < B_3 < 160.5 \ \& \ 85.8 < B_4 < 131.4 \ \& \ 137.7 < B_5 < 238.1 \ \& \ 64.2 < B_7 < 126.4$ Then class=Developing land </p>

表 3 番禺实验区粒子群(PSO)的分类精度评价结果

实际	分类									使用精度/%
	建成区	山体	水体	果园	农田	休耕地	基塘	建设用地	总和	
建成区	288	1	2	1	8	19	3	7	329	87.6
山体	1	98	1	9	5	0	0	1	115	85.2
水体	1	3	313	2	2	1	28	1	351	89.2
果园	9	10	3	184	35	4	2	0	247	74.5
农田	1	4	2	31	259	7	1	2	307	84.4
休耕地	10	1	1	6	4	107	1	1	131	81.3
基塘	11	3	37	1	0	4	285	1	342	83.4
建设用地	10	1	0	3	2	3	1	158	178	88.7
总和	331	121	359	237	315	145	321	171	2000	
生产精度/%	87.0	81.0	87.2	77.6	82.2	73.8	88.7	92.4		

总精度=84.6% Kappa 系数=0.821

表 4 番禺实验区 See5.0 决策树的分类精度评价结果

实际	分类									使用精度/%
	建成区	山体	水体	果园	农田	休耕地	基塘	建设用地	总和	
建成区	281	2	3	0	9	22	5	7	329	85.4
山体	1	96	0	8	8	0	1	1	115	83.4
水体	2	4	304	3	3	1	33	1	351	86.6
果园	7	12	2	183	36	6	1	0	247	74.1
农田	1	5	2	31	258	8	1	1	307	84.0
休耕地	13	2	0	5	5	104	1	1	131	79.4
基塘	14	2	60	3	1	6	254	2	342	74.3
建设用地	13	1	1	5	2	1	0	155	178	87.1
总和	332	124	372	238	322	148	296	168	2000	
生产精度/%	84.6	77.4	81.7	76.9	80.1	70.3	85.8	92.3		

总精度=81.8% Kappa 系数=0.788

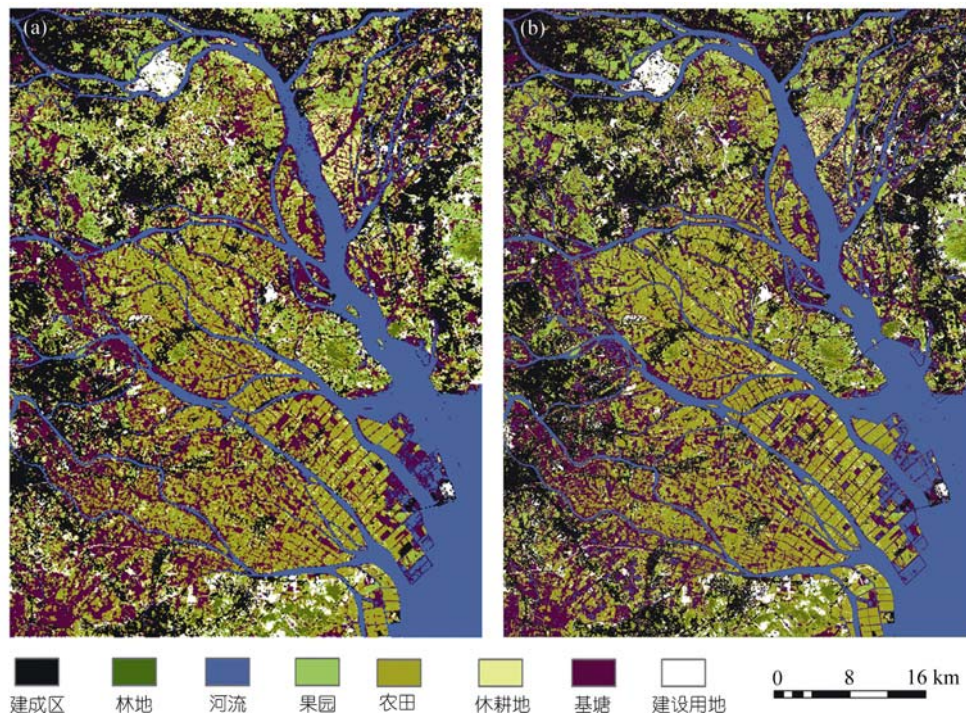


图 5 番禺实验区土地利用分类结果

(a) PSO 分类结果; (b) See5.0 分类结果

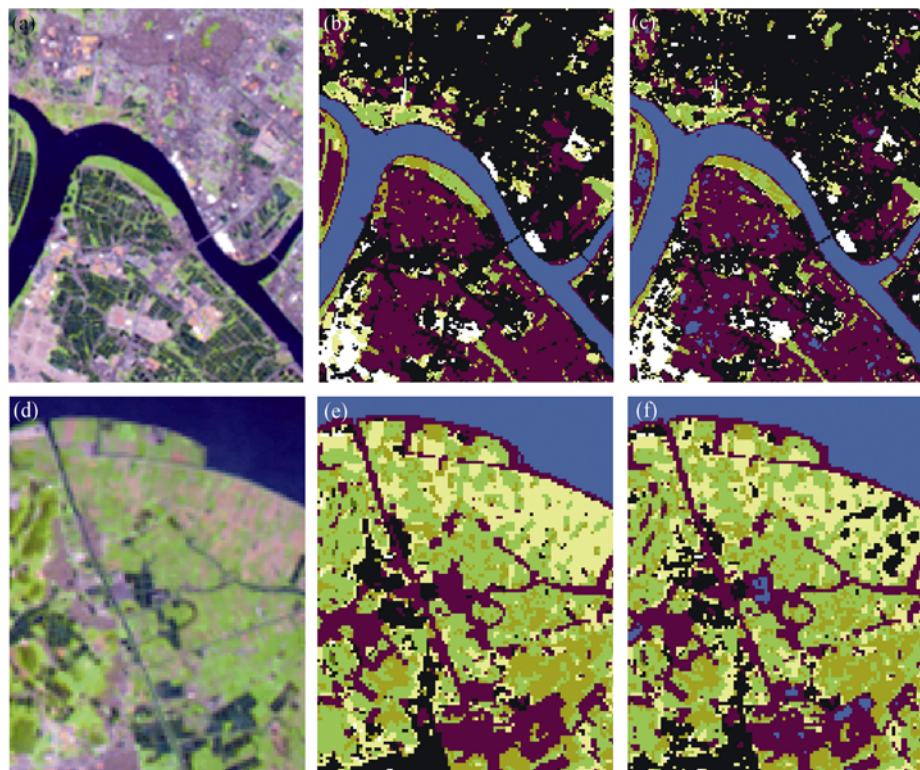


图 6 局部放大地区的土地利用分类结果

(a) TM image; (b) PSO 分类结果; (c) See 5.0 分类结果; (d) TM image; (e) PSO 分类结果; (f) See 5.0 分类结果

$$\text{Kappa} = \frac{M \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{M^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}, \quad (7)$$

式中, x_{ii} 表示错误矩阵主对角线上的元素, x_{i+} 表示错误矩阵 i 行的和, x_{+i} 表示错误矩阵 i 列的和. 根据公式 (7), 分别计算了两种分类方法的 Kappa 系数(表 3, 4), 其中, 粒子群遥感分类的 Kappa 系数为 0.821, See5.0 决策树遥感分类的 Kappa 系数为 0.788, 两者之间的精度相差明显. 可见, 粒子群(PSO)分类的效果比 See5.0 决策树方法的效果要好.

4 结论

智能式遥感分类是遥感研究的新热点之一. 当地表状况较复杂时, 目前的遥感分类方法对于线性不可分的地物往往难以区分, 有必要引进智能式的方法来提高分类精度. 为此, 本文提出了一种基于粒子群的智能式遥感分类新方法. PSO 算法实际上是一种多智能体算法, 简单的智能个体通过相互协作和竞争而表现出复杂的群集智能行为, 不会由于一个或者某几个智能个体的故障而影响整个问题的求解; 与进化算法相比较, PSO 采用简单的速度-位移模型, 避免了复杂的遗传操作, 同时它的记忆能力使其可以跟踪当前的搜索情况并动态调整搜索策略, 具有较强的全局收敛能力和鲁棒性, 因此, 非常适合地理复杂规律的提取.

文章首次将基于粒子群的分类规则挖掘算法 (PSO-Miner) 尝试性地引进到遥感影像中, PSO-Miner 算法模仿鸟类觅食的方式构造分类规则, 通过序列覆盖算法移除规则所覆盖的训练数据, 直至所有的规则挖掘完毕. 本文设计的 PSO-Miner 算法, 粒子能自动寻找每个波段的最优分割点. 并且该算法所提取的 If-Then 分类规则, 能更方便和准确地描述自然界中的复杂关系, 比数学公式更容易让人理解.

将该算法应用于番禺的遥感影像分类中, 取得了较好的分类结果. 并与 See5.0 决策树方法进行了对比研究, 其中 PSO 分类方法的总精度为 84.6%, Kappa 系数为 0.821, 而 See5.0 决策树方法的总精度为 81.8%, Kappa 系数为 0.788. 这表明粒子群(PSO)的分类精度比 See5.0 决策树方法更高.

参 考 文 献

- 1 Wilkinsom G G. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE T Geosci Remote Sens*, 2005, 43(3): 433—440[DOI]
- 2 Strahler A H. The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sens Environ*, 1980, 10: 135—163[DOI]
- 3 Aksoy S, Koperski K, Tusk C, et al. Learning bayesian classifiers for scene classification with a visual grammar. *IEEE T Geosci Remote Sens*, 2005, 43(3): 581—589[DOI]
- 4 Wilkinsom G G. A review of Current issue in the integration of GIS and Remote sensing data. *Int J Geogr Inf Sci*, 1996, 10(1): 85—101[DOI]
- 5 Pal M, Mather P M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens Environ*, 2003, 86: 554—565[DOI]
- 6 Friedl M A, Brodley C E. Decision tree classification of land cover from remotely sensed data. *Remote Sens Environ*, 1997, 61: 399—409[DOI]
- 7 Huang C, Davis L S, Townshend J R G. An assessment of support vector machines for land cover classification. *Int J Remote Sens*, 2002, 23: 725—749[DOI]
- 8 Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE T Geosci Remote Sens*, 2004, 42: 1778—1790[DOI]
- 9 Frizzelle B G, Moody A. Mapping continuous distributions of land cover: a comparison of maximum-likelihood estimation and artificial neural networks. *Photogramm Eng Remote Sens*, 2001, 67(6): 693—705
- 10 Heerman P D, Khazenie N. Classification of multispectral remote sensing data using a back propagation neural network. *IEEE T Geosci Remote Sens*, 1992, 30(1): 81—88[DOI]
- 11 Wang F. Fuzzy supervised classification of remote sensing images. *IEEE T Geosci Remote Sens*, 1990, 28(2): 194—201[DOI]
- 12 Bandyopadhyay S, Pal S. Pixel classification using variable string genetic algorithms with chromosome differentiation. *IEEE T Geosci Remote Sens*, 2001, 39(2): 303—308[DOI]
- 13 Foody G M, Mathur A. A relative evaluation of multi class image classification by support vector machines. *IEEE T Geosci Remote Sens*, 2004, 42(6): 1335—1343[DOI]
- 14 Machado Th R, Lopes H S. A Hybrid Particle Swarm Optimization Model for the Traveling Salesman Problem. *Adaptive and Natural Computing Algorithms*. New York: Springer, 2005. 255—258
- 15 Kennedy J. Stereotyping: improving particle swarm performance with cluster analysis. *Proceedings of the Congress on Evolutionary Computing*. Piscataway, NJ: IEEE Service Center, 2000. 1507—1512
- 16 Ting T O, Rao M V C, Loo C K, et al. Solving unit commitment problem using hybrid particle swarm optimization. *J Heuristics*,

- 2003, 9 (6): 507—520[DOI]
- 17 Kwang M S, Weng H S. Multiple ant-colony optimization for network routing. Proceedings of the First International Symposium on Cyber Worlds, 2002. 277—281
- 18 Omran M. Particle Swarm optimization methods for pattern Recognition and Image Processing. Pissertation for the Doctoral Degrel. University of Pretoria, 2005
- 19 Kennedy J, Eberhart R C. Particle Swarm Optimization. Proceedings of IEEE, International Conference on Neural Networks, Piscataway, NJ, 1995. 1942—1948
- 20 Kennedy J, Eberhart R C, Shi Y. Swarm Intelligence. San Francisco: Morgan Kaufman Publishers, 2001
- 21 Clerc M, Kennedy J. The particle swarm-explosion, stability and convergence in a multidimensional complex space. IEEE Trans Evolut Comput, 2002, 6(1): 58—73[DOI]
- 22 Sousa T, Neves A, Silva A. A particle swarm data miner. In: 11th Portuguese Conference on Artificial Intelligence, Workshop on Artificial Life and Evolutionary Algorithms, 2003. 43—53
- 23 Liu Y, Qin Z, Shi Z W, et al. Rule discovery with particle swarm optimization. Lect Notes Comput Sci, 2004, 291—296
- 24 Shi Y H, Eberhart R C. Parameter selection in particle swarm optimization. The Seventh Annual Conference on Evolutionary Programming, Washington DC, 1998. 591—600
- 25 Congalton R G. A review of assessing the accuracy of classification of remotely sensed data. Remote Sens Environ, 1991, 37: 35—46[DOI]