

文章编号: 1672-8785(2011)08-0018-07

基于智能生态细胞的遥感信息提取方法

蒋样明¹ 崔伟宏¹ 董前林²

(1. 中国科学院遥感应用研究所, 北京 100101;

2. 中国矿业大学, 北京 100083)

摘 要: 提出了一种基于智能生态细胞的遥感信息提取方法。该方法综合了作物的物候特征、光谱演化特征和空间相关性的反馈信息, 融合了元胞自动机、模糊推理和 BP 神经网络智能计算思想。以 2009 年多时相 Landsat7 ETM 遥感影像作为数据源, 进行了基于智能生态细胞的烤烟种植面积监测。监测结果表明, 该方法可以有效提高遥感信息提取精度, 具有实际应用价值。

关键词: 智能生态细胞; 元胞自动机; 模糊推理; BP 神经网络; 物候特征

中图分类号: TP79 **文献标识码:** A **DOI:** 10.3969/j.issn.1672-8785.2011.08.004

Remote Sensing Information Extraction Method Based on Smart Ecological Cells

JIANG Yang-ming¹, CUI Wei-hong¹, DONG Qian-lin²

(1. Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China;

2. China University of Mining and Technology, Beijing 100083, China)

Abstract: A remote sensing information extraction method based on smart ecological cells is put forward. In the method, the crop phenological characteristics, spectral evolution characteristics and special correlation feedback are combined and the cellular automata, fuzzy reasoning and BP neural network intelligent computation are fused. By taking the multi-temporal Landsat-7 ETM images as the experimental data, the monitoring of tobacco acreage planting area based on smart ecological cells is carried out. The monitoring result shows that this method can effectively improve the extraction accuracy of remote sensing information and is of value to practical applications.

Key words: smart ecological cell; cellular automata; fuzzy reasoning; BP neural network; phenological characteristics

0 引言

目前, 遥感图像已经成为地球观测的重要信息源。遥感信息的提取与分类是遥感图像分析中的重要研究内容。遥感图像是某一时刻下地表生态环境和生态特征的综合反映, 它具有

较强的时效性。由于受传感器空间分辨率的限制, 遥感图像中的像元是地表生态环境、地表生态特征和大气环境的综合表征。遥感图像的这种特征不仅会影响地物识别和分类精度, 而且还是遥感定向、定量化发展的瓶颈之一^[1]。

针对以上情况, 人们分别提出了监督分类

收稿日期: 2011-06-08

基金项目: 中国博士后科学基金(20080430586); 国家科技支撑项目(2008BAK50B01)

作者简介: 蒋样明(1983-), 男, 浙江杭州人, 副研究员, 博士, 主要研究方向为地图学与地理信息系统。

E-mail: jym_wm@126.com

法、非监督分类法、线性混合光谱分解法、模糊监督分类法、神经网络模型分解法^[2]、自动形态学端元提取方法^[3-4]、模糊分类器模型^[5,6]、高斯混合判别分析模型^[7]以及支持向量机模型(Support Vector Machine, SVM)^[8]等。他们在利用这些模型和方法提取遥感信息时,只考虑了地物的光谱特征,而忽略了地物周边生态环境的影响^[9]。本文提出一种基于智能生态细胞的遥感信息提取方法,该方法在考虑地物光谱特征的同时,充分利用了遥感影像各像元的周边生态环境信息,可提高遥感影像识别、分类和信息提取的精度。

1 基于智能生态细胞的遥感信息提取方法

1.1 智能生态细胞的基本概念

智能生态细胞(Smart Ecological Cell, SEC)高度抽象和概括了对象的空间特征、时间特征以及环境特征。每个智能生态细胞包含了地表对象的生态特征、生态环境信息和生态过程信息。规则的、离散的、有限状态下的智能生态细胞按照一定规则,组成在离散时间维上演化的、非线性的局部生态环境动力学系统。其中,每个生态细胞与有限的其他生态细胞进行信息交换,构成具有几何一致性的邻域。上述思想基于局部到整体、自下而上的研究思路,认为局部生态环境系统的状态由所有生态细胞相互作用而形成,每个生态细胞的特征及其下一时刻的变化由其初始状态、自身生态特征和邻域生态细胞对其的作用而决定。在决定任何一个生态细胞的行为时,可充分利用系统要素间的反馈信息^[1]来提高生态细胞特征和状态的准确度^[10],并利用人工神经网络(Artificial Neural Network, ANN)来描述该生态环境动力学系统的复杂非线性关系,然后将每个生态细胞看作分布储存于神经网络内的各神经元^[11-12]。另外,为了使模型和算法能够适应复杂、多变、模糊的外部环境作用,可在生态细胞中引入模糊集理论、模糊逻辑和模糊推理功能。因此,智能生态细胞具备了自学

习、自适应以及并行处理定量/定性知识和模糊信息的功能。

1.2 智能生态细胞的数学模型

智能生态细胞由生态细胞空间、生态细胞状态、邻域以及转换规则组成。在数学形式上,任意一个智能生态细胞均可表达为一个四元组 $SEC=(Z^n, S, N, f)$ 。其中, Z^n 表示智能生态细胞的空间结构,属于 n 维欧氏空间; S 为智能生态细胞的有限状态集合, $S=\{S_1, S_2, S_3, \dots, S_i, \dots, S_k\}$, S_i 为智能生态细胞的第 i 个状态; N 为中心智能生态细胞的邻域, $N=(x_1, x_2, \dots, x_i, \dots, x_n)$, x_i 为相邻智能生态细胞相对于中心智能生态细胞的位置; f 为 $S_i^t \rightarrow S_i^{t+1}$ 的局部转换规则,即为生态细胞状态从 t 时刻进化到 $t+1$ 时刻所遵循的规则。

智能生态细胞是局部生态系统最基本的组成部分。每个智能生态细胞都可以采用一维、二维或多维欧几里德空间结构表达,而局部生态系统的空间则由智能生态细胞排列成的 n 维阵列表达。智能生态细胞的有限状态集合代表了生态细胞的典型状态。在获取智能生态细胞的状态时,通过引入模糊集理论和揉和人工智能的其他手段,可以满足智能生态细胞的差异性和多样性特征。模糊计算可以处理不精确的模糊输入信息,有效降低灵敏度和精确度要求,从而模拟出客观世界的非精确、非线性的信息处理能力。通过将人工神经网络的学习机制和模糊逻辑的推理过程相结合,可以使生态细胞具有状态自动探测功能。

1.3 基于智能生态细胞的遥感信息提取模型设计

基于智能生态细胞的遥感信息提取模型可以充分利用像元的空间结构信息及邻域的光谱特征。下面介绍该模型的具体设计信息。

1.3.1 智能生态细胞空间

智能生态细胞空间的欧几里德空间采用方格型结构,相当于遥感影像的像素灰度值矩阵。因此可以将智能生态细胞理解成一个个离散的

像素点。

1.3.2 智能生态细胞的状态

根据遥感信息提取对象的生态过程特征,选取典型时期遥感影像的像素灰度值作为智能生态细胞的状态。为了便于智能辨识的设计和实现,对智能生态细胞的输入、输出值进行规范化,使得输入、输出值处在规定的范围内,同时将输入值转化为模糊量。图2为智能生态细胞信息提取的流程图。其中,模糊逻辑推理将决定输出量的一个分布函数;清晰化过程就是将输出量的分布函数转换成规范化的输出量;最后系统将规范化的输出量转换为实际的输出值。

1.3.3 智能生态细胞的邻域

智能生态细胞的邻域可以是冯-诺依曼模型、摩尔型或马哥勒斯型结构等。本文采用摩尔型邻域结构,其定义如下:

$$N_{Moore} = \{v_i = (v_{ix}, v_{iy}) | |v_{ix} - v_{0x}| \leq n, |v_{iy} - v_{0y}| \leq n, (v_{ix}, v_{iy}) \in Z^2\} \quad (1)$$

式中, v_{ix} 和 v_{iy} 为邻居智能生态细胞的行、列坐标值, v_{0x} 为中心智能生态细胞的行、列坐标值, n 为正整数。

1.3.4 智能生态细胞的转换规则

模型设计的关键在于转换规则的设计,而转换规则又是基于智能生态细胞的模型中最复杂的部分。本文设计的转换规则融合了模糊控制和人工神经网络。在智能生态细胞信息提取过程中,我们首先对遥感影像的像素值进行规

范化,并通过模糊化处理将像素值转化为模糊量;然后定义神经网络结构,选择经过模糊化处理的训练样本数据集,此时神经网络会根据样本进行自学习和自我修正偏差,并根据样本数据自行“总结规则”;在神经网络的输出经过模糊系统处理后,根据相应的模糊规则对其进行模糊推理;最后对推理结果进行清晰化和规范化,输出遥感信息提取结果。

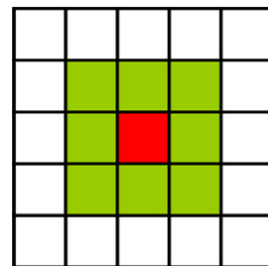


图1 N=1 摩尔型结构

将模糊论域划分为4个等级(ZO: 概率为零; PS: 小概率; PM: 中概率; PB: 大概率); 隶属函数采用三角形函数(见式(2))。图3为隶属函数的图像。模糊规则采用 *if... then...* 格式来表达, 如 $R_i: \text{if } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \cdots \text{and } x_n \text{ is } A_{in} \text{ then } y = f_1(x_1, x_2, \cdots, x_n)$ 。

$$u(x) = \begin{cases} \frac{x-b}{a-b}, & b \leq x \leq a \\ \frac{c-x}{c-a}, & a < x \leq c \\ 0, & x < b \text{ 或 } x > c \end{cases} \quad (2)$$

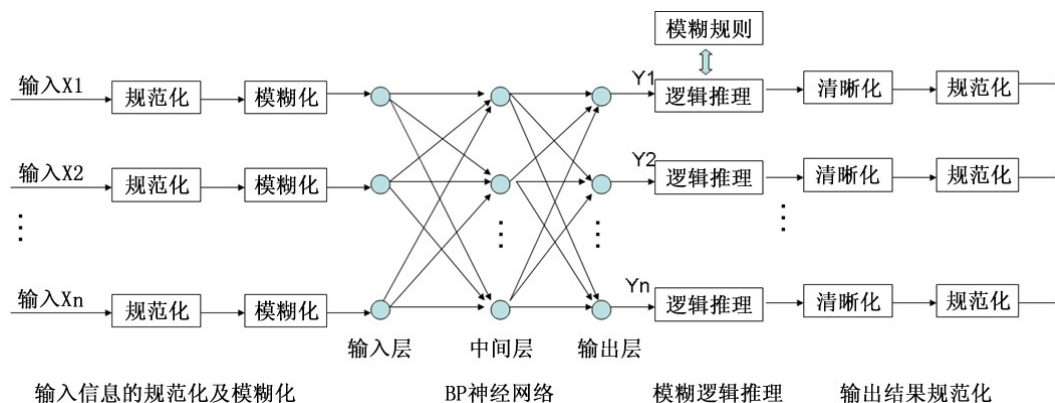


图2 智能生态细胞信息提取的流程图

利用 BP 神经网络计算来自动学习推理规则, 其具体步骤如下:

(1) 初始化神经网络的各层权值 W_{ij} 和神经元阈值 θ_j 以随机任意小值。

(2) 给定训练数据集, 即提供输入向量 X 和期望输出 $\bar{y}$ 。

(3) 计算实际输出 y (见式(3))。式中, y_j 为输出层第 j 个单元的输出, W_{ij} 为第 i 个单元与第 j 个单元的连接权值。

$$y_j = g(s_j) = \frac{1}{1 + \exp\left(-\sum_{i=1}^n W_{ij} * x_i + \theta_j\right)} \quad (j = 1, 2, \dots, q) \quad (3)$$

(4) 计算网络输出误差 $E = \sum (\bar{y}_j - y_j)^2 / 2$, 其中 $\bar{y}_j$ 与 y_j 为输出层第 j 个单元的期望输出与实际输出。

(5) 调整权值, 按误差反向传播方向修正权值 $W_{ij}(t+1) = W_{ij}(t) + \eta \delta_j y_i$ 。其中, η 为大于零的增益, δ_j 为 j 单元的网络误差。

(6) 判断误差是否满足要求。若满足, 则训练结束, 否则返回第 (3) 步重复进行计算。

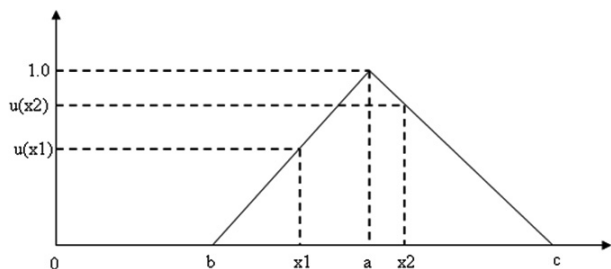


图 3 隶属函数的图像

2 基于智能生态细胞提取烟田空间分布信息

2.1 物候、光谱特征及生态细胞状态

选择云南省某县作为实验区。实验区内烤烟的物候特征为: 3 月初至 4 月中旬为烟苗培育期; 4 月下旬至 5 月中旬为烟苗移栽期; 5 月下旬至 6 月上旬为伸根期; 6 月中旬开始进入旺长期; 7 月下旬至 8 月中旬是烤烟的成熟期。其种植特点是在烟苗培育期进行烟田整平与起垄, 在移栽时覆盖地膜。春玉米具有相似的种植特点, 但春玉米的播种时间为每年的 3 月下旬至 4 月上旬, 一般采用先播种后覆膜的方式, 于 8 月底 9 月初成熟。与烤烟相比, 其地膜覆盖时间存在近一个月的时间差。

图 4 所示为 5 月下旬烤烟移栽后试验区内主要地物的光谱曲线。从图 4 中可以看出, 地膜的光谱特征与其他地物存在明显的差异(如曲线 b), 这有利于辅助识别烟田地块。烟田和玉米的植被覆盖度^[13-14]及地膜覆盖度会随着时间的推移而逐渐演变。因此, 根据其演变规律选择适宜的遥感监测时间, 成为区分烟田和玉米的关键。根据试验区内烤烟和春玉米的生长物候期, 通过模拟得到了其植被覆盖度和地膜可视度的演变趋势曲线, 如图 5 所示。图 6 所示为烤烟移栽期和旺长期时烤烟和玉米的 TM 光谱曲线。从图 6 中可以看出, 6 月初烤烟和玉米的 TM 光谱曲线存在比较明显的差异(曲线 a 和 b)。此时在 TM 影像中, 烟田表现为土地、地膜和烟苗的混合光谱, 而玉米地则主要表现为植被的光谱

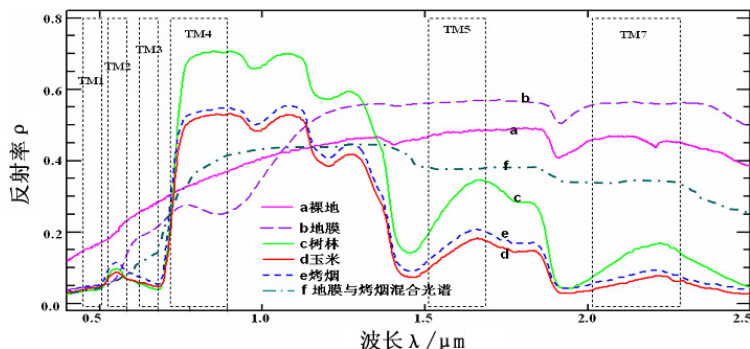


图 4 烟草遥感监测实验区内主要地物的光谱曲线

特征; 7 月初烤烟和玉米的植被覆盖度都很高, 地膜可视度接近于 0, 主要表现为烤烟和玉米的光谱特征 (曲线 c 和 d)。

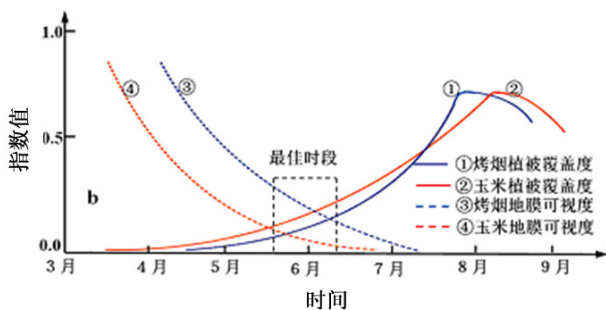


图 5 植被覆盖度和地膜可视度的示意图及其演变趋势

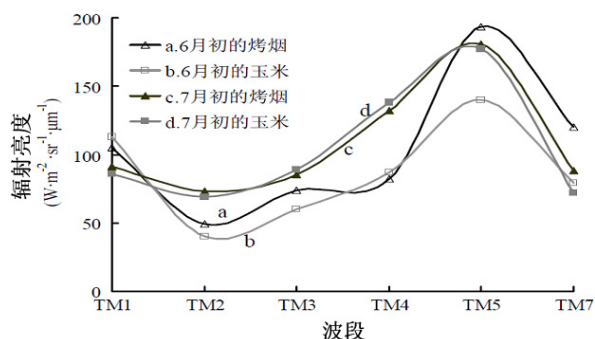


图 6 烟草遥感监测实验区内不同时期地物的 TM 光谱曲线

根据烤烟的物候特征, 选择烟苗培育期、移栽期和旺长期 3 个时期的 Landsat7 ETM 遥感影像作为输入数据源, 并将育苗期内整平、起垄的裸地光谱特征, 移栽期内烟苗和地膜的混合光谱特征以及旺长期内烤烟的光谱特征作为智能烤烟生态细胞的状态。将模糊论域划分为 4 个等级 (ZO: 概率为零; PS: 小概率; PM: 中概率; PB: 大概率)。隶属函数采用三角形函数。将连续的三角形隶属函数离散化, 形成离散域的一组向量, 由此构成育苗期裸地 {Bi}、移栽期地膜 {Fi} 和旺长期烤烟光谱特征 {Ti} 的 3 组模糊集。

2.2 烤烟生态细胞转换规则的设计

模糊规则以 If ... Then ... 格式出现。根据 {Bi}、{Fi}、{Ti} 模糊集, 模糊规则的定义如表 1 所示。

模糊推理采用 Matlab 中的 Fuzzy 工具进行模糊计算, 具体操作如下:

- (1) 应用 *newfis()* 函数创建一个 FIS(Fuzzy Inference System) 对象;
- (2) 应用 *addvar()* 函数增加模糊变量;
- (3) 应用 *addmf()* 函数创建模糊集合;
- (4) 通过 *addrule()* 函数增加模糊推理规则, 即将表 1 中的模糊规则添加到模糊计算中;
- (5) 应用 *evalfis()* 函数进行模糊推理, 其表达式为 $output = evalfis(input, fismat)$, 其中 *fismat* 为 FIS 对象。

表 1 模糊规则的定义

If		Then	
地膜	裸地	烤烟	烟田
ZO	ZO	ZO	ZO
PS	ZO	ZO	PS
PM	ZO	ZO	PM
PB	ZO	ZO	PB
ZO	PS	ZO	ZO
PS	PS	ZO	PS
...	...	...	...
PB	PB	PB	PB

2.3 烤烟生态细胞 BP 神经网络的设计

采用 Matlab 中的 *newff()* 函数建立神经网络; 采用 *initff()* 函数对神经网络进行初始化; 采用 *train()* 函数训练前向网络; 采用 *sim()* 函数完成前向网络的计算。通过 Matlab 中的函数实现神经网络训练, 其中输入数据为育苗期裸地、移栽期地膜、旺长期烤烟的光谱信号值, 由这些输入值组成矩阵 P 和 T。使用 *newff()* 函数创建一个包含 3 个输入节点、8 个中间层节点和 3 个输出节点的三层网络模型。其中, 各层间的传输函数采用 *Sigmoid* 函数, 其表达式为

$$net=newff(minmax(P),\{3,8,3\},$$

$$\{'logsig','logsig','logsig'\},'trainlm') \quad (4)$$

然后对神经网络进行初始化:

$$[w,b1,w2,b2,w3,b3]=initff(P,3,'logsig',$$

$$8,'logsig',3,'logsig') \quad (5)$$

接着定义网络训练参数并训练网络:

```
net.trainParam.goal=0.00001;
net.trainParam.epochs=20000;
net.trainParam.show=20;
net.trainParam.lr=0.5;
[net,tr]=train{net,P,T}
```

最后由经过训练的网络计算出网络的实际输出值 $A=\text{sim}(\text{net},P)$ 。

2.4 智能生态细胞精确化规则

(1) 初始化: 将智能生态细胞划分为 3×3 的亚像元, 并根据各细胞中裸地、烟田和玉米地所占的比例分配亚像元个数, 从而重新构成精度更高的生态细胞系统。

(2) 设置损失交换概率 m ($m=0.05$)。

(3) 进行系统转化, 其中每一步转化由 k (k 的大小一般由地物种类和亚像元个数决定) 步子过程组成。每一个子过程包括以下步骤:

第一步: 随机选择一个邻居状态与其本身状态不完全相同的生态细胞。

第二步: 在生态细胞空间中随机选择另外一个生态细胞。如果该生态元胞的状态与原始生态细胞的相同, 或者该生态细胞的所有邻居状态都与其相同, 那么结束子过程, 否则向下执行第三步。

第三步: 计算“交换效益” G (G 表示在交换两个生态细胞之后, 新状态与生态细胞状态

相同的邻居生态细胞个数之和减去交换之前状态与生态细胞状态相同的邻居生态细胞个数)。 $G > 0$ 表示交换两个生态细胞可以提高整体上的空间相关性, 此时交换两个元胞; 否则以损失交换概率 m 来交换两个生态细胞。

第四步: 结束子过程。

2.5 基于智能生态细胞的烟田空间分布信息提取流程及实验结果

(1) 应用实验区内烟田土地整平、起垄阶段中 2009 年 4 月 17 日的 Landsat7 ETM 数据, 计算该时相各生态细胞的植被指数, 并根据裸地的光谱特征计算各生态细胞裸地的概率。应用移栽期阶段中 2009 年 5 月 11 日的 Landsat7 ETM 数据, 并根据地膜的光谱特征计算该时相各生态细胞覆膜的概率。应用旺长期内 2009 年 6 月 20 日的 Landsat7 ETM 数据, 并根据该时相烤烟和地膜的混合光谱特征、玉米和地膜的混合光谱特征计算该时相各生态细胞玉米的概率和烤烟的概率。

(2) 利用模糊论域和三角形隶属函数将各生态细胞的输入值转化为 3 组模糊集 $\{B_i\}$ 、 $\{F_i\}$ 、 $\{T_i\}$ 。

(3) 将模糊集中的值作为 BP 神经网络的输入值, 然后利用前面设计的 BP 神经网络进行计算。

(4) 对神经网络的输出进行模糊函数处理,

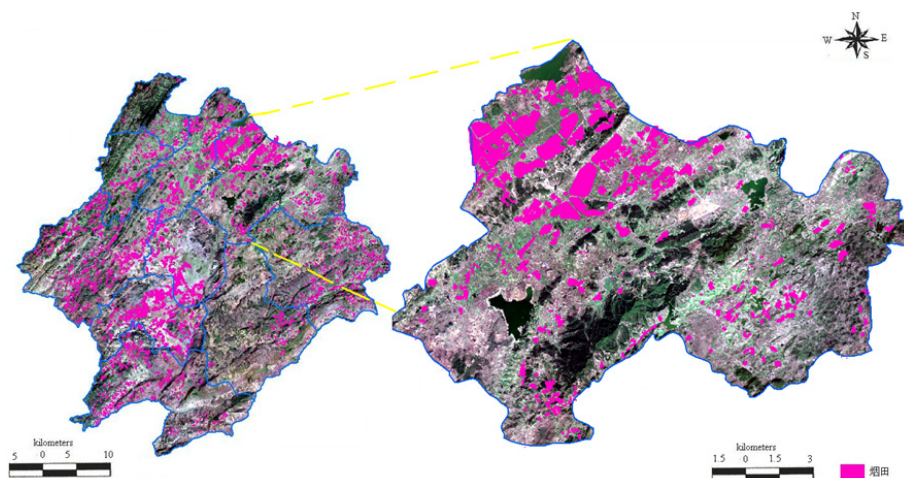


图 7 实验区内烟田面积的空间分布图

表 2 实验区内 2009 年的烤烟种植面积遥感监测结果的统计表

区域	白*镇	金*镇	旧*镇	河*乡	塘*乡	舞*镇	向*乡	水*乡	中*镇	逸*镇	总计
种植面积 / 亩	36300	18000	6700	3100	4400	22800	3000	12000	16000	24000	146300
A 监测面积 / 亩	39975	19566	7262	3394	4800	24829	3309	13104	17408	26328	159975
A 偏差率 / %	9.3	8.7	8.4	9.5	9.1	8.9	10.3	9.2	8.8	9.7	9.3
A 实测采样	GPS 采样地块数 / 块			837	监测符合地块数 / 块			763	偏差率 / %		8.9
	GPS 测量面积 / 亩			20247	监测面积 / 亩			22130	偏差率 / %		9.3
B 监测面积 / 亩	38737	19150	7252	3325	4727	24010	3295	12681	17439	25932	156548
B 偏差率 / %	6.7	6.4	8.2	7.3	7.4	5.3	9.8	5.7	9.0	8.1	7.0
B 实测采样	GPS 采样地块数 / 块			837	监测符合地块数 / 块			792	偏差率 / %		5.4
	GPS 测量面积 / 亩			20247	监测面积 / 亩			21542	偏差率 / %		6.4

注: A 表示课题组采用基于多时相 Landsat7 ETM 遥感数据的监督分类、多源数据融合、目视解译综合的烤烟面积监测方法; B 表示采用基于智能生态细胞的烟田空间分布信息提取方法进行烤烟面积遥感监测; 若实测烟田中心与遥感监测烟田中心的距离在阈值(100 m)以内, 则认为遥感监测地块位置准确。

然后利用前面定义的模糊规则和 Matlab 模糊推理方法计算各生态细胞烤烟种植的概率。

(5) 利用智能生态细胞精确化规则对智能生态细胞中裸地、烟田和玉米地的空间位置及其所占比重进行亚像元级定量描述。

采用基于智能生态细胞的烟田空间分布信息提取方法进行烤烟面积遥感监测, 获得了实验区内 2009 年烤烟种植面积的遥感监测结果。图 7 为烟田面积的空间分布图。利用各个地区的烟叶种植合同数据和应用 GPS 手持机实地采样的烟田样本对监测结果进行了验证, 并且在 ARCGIS 软件中将实测采样数据与监测结果进行了 Erase 分析(统计结果见表 2)。

3 结果与讨论

本文提出一种基于智能生态细胞的遥感信息提取方法, 即根据实验区内典型作物的物候特征和典型的生态特征选取合理的多时相遥感影像作为输入数据源, 综合利用智能生态细胞的模糊推理、BP 神经网络智能计算和生态细胞空间相关性的反馈信息, 提高了实验区内烤烟种植面积的监测精度。通过将遥感监测结果与实验区烟叶种植合同数据进行对比可知, 烟田遥感监测的平均精度偏差约为 7.0 %; 通过与实测采样数据进行对比可知, 烟田遥感监测的位置精度偏差为 5.4 %, 面积精度偏差为 6.4 %。

通过将基于智能生态细胞的烟田面积遥感监测结果与基于多时相遥感数据的监督分类、多源数据融合、目视解译综合的烤烟面积监测方法的监测结果进行对比可知, 基于智能生态细胞的遥感信息提取方法有效地提高了烤烟种植面积的监测精度, 具有实际应用价值。

参考文献

- [1] 王旭红, 郭建明, 贾百俊, 等. 元胞自动机的遥感影像混合像元分类[J]. 测绘学报, 2008, 37(1): 42-48.
- [2] Chang L C, Shao S S, Simth J A, et al. Linear Spectral Random Mixture Analysis for Hyperspectral Imagery[J]. *IEEE Transactions on Geosciences and Remote Sensing*, 2002, 40(2): 375-392.
- [3] Plaza A, Chen L C. Recent Advances in Hyper spectral Signal and Image Processing [M]. India: Transworld Research Net-work, 2006.
- [4] Plaza A, Martinez P, Perez R, et al. A Quantitative and Comparative Analysis of End Member Extraction Algorithms from Hyper Spectral Data [J]. *IEEE Transactions on Geosciences and Remote Sensing*, 2004, 42(3): 650-663.
- [5] Foody G M. Sharpening Fuzzy Classification Output to Refine The Representation of Sub-pixel Land Cover Distribution [J]. *International Journal of Remote Sensing*, 1998, 13(2): 2593-2599.
- [6] Foody G M. Approaches for The Production and Evaluation of Fuzzy Land Cover Classifications from Remotely-sensed Data [J]. *International Journal of Remote Sensing*, 1996, 17(7): 1317-1340.

(下转第 28 页)