

文章编号:1673-9469(2020)01-0103-06

DOI: 10.3969/j.issn.1673-9469.2020.01.017

基于 GF-1 和 MODIS 时序 NDVI 的种植结构提取

李家宝,王贺封,张安兵*,李红红

(河北工程大学 矿业与测绘工程学院,河北 邯郸 056038)

摘要: 为提高作物识别精度、扩展数据适用范围,以两幅 16 m 空间分辨率的高分一号(GF-1)数据和 23 幅 250 m 分辨率的 MODIS 数据为数据源,首先进行 GF-1 的监督分类和 MODIS NDVI 时间序列的构建,然后利用监督分类结果和时间序列数据共同构造决策树模型,对石津灌区的主要作物进行遥感解译。由地面样本点和随机点进行精度检验,总体提取精度达到 93.13%,较单纯的监督分类,解译精度提高 10% 以上。结果表明该方法不仅能较好识别作物时空分布信息,而且能更好地发挥数据各自的优势。

关键词: GF-1; MODIS; 监督分类; 决策树模型; 种植结构

中图分类号: P237

文献标识码: A

Extraction of Crop Planting Structure Based on GF-1 and MODIS Sequential NDVI Characteristics

LI Jiabao, WANG Hefeng, ZHANG Anbing*, LI Honghong

(School of Mining and Mapping Engineering, Hebei University of Engineering, Handan, Hebei 056038, China)

Abstract: In order to improve the accuracy of crop identification and expand the scope of data application, two pieces of GF-1 data with 16 m spatial resolution and 23 pieces of MODIS data with 250 m resolution were used as data sources. First, the supervised classification of GF-1 and the construction of MODIS NDVI time series were carried out. Then, the decision tree model was constructed by using the supervised classification results and time series data characteristics to identify the main crops in Shijin Irrigation Area. The overall accuracy of extraction reaches 93.13%, which is more than 10% higher than the result of supervised classification. The results show that the method can not only recognize the spatial and temporal distribution information of crops, but also give full play to the advantages of the data.

Key words: GF-1; MODIS; supervised classification; decision tree model; crop planting structure

作物的种植结构是反映作物类型、空间分布、种植面积和轮作模式的综合农业信息,其时空变化信息是农业可持续发展的重要研究内容^[1]。逐层统计汇总是获得农作物种植结构信息的传统方式,但存在实地大面积调查测量误差大、耗时耗力和时间滞后的问题^[2]。遥感技术因其高时效、宽范围和低成本的优势,近年来在农业、水资源管理方面的应用越来越普及,为大区域准确和快速获取农作物种植结构提供了新的技术手段^[3]。

作物在遥感影像上呈现的不同光谱特征、时

相特征和空间特征是种植结构遥感解译的理论依据^[4]。单一影像的解译方法主要是根据作物光谱差异并结合作物生长的关键期,找到识别作物的最佳时相,操作简单;但由于受“异物同谱”现象影响,为确保作物识别精度,要求影像具有较高空间分辨率和较好的影像质量^[5-6]。而时间序列影像的解译方法充分利用作物整个生长生育阶段的信息,一定程度上解决了单一影像方法的混合像元问题。归一化植被指数 NDVI (normalized difference vegetation index) 反映了植被生长状态及植被

收稿日期:2019-11-11

基金项目:河北省自然科学基金资助项目(D2015402134, D2017402159)

作者简介:李家宝(1992-),女,河南新乡人,硕士研究生,从事农业遥感方面的研究。

* 通讯作者:张安兵(1978-),男,山东郓城人,博士,教授,研究方向为 GPS 数据处理、变形监测、高光谱遥感应用等。

覆盖度,NDVI的时序变化对应了植被的生长和衰落等活动过程^[7],并反映出农作物各自的生长差异,因此NDVI被广泛应用于植被信息提取和地物识别^[8-9]。但应注意到,由于数据源的限制,目前能够构成完整作物生长周期的时间序列数据大多集中在中低空间分辨率,难以获得较好的提取精度。为此,本文探讨运用GF-1监督分类结果和MODIS NDVI时间序列共同参与决策树构建来提取区域作物种植结构的方法,克服单一数据源或单一提取方法缺点的同时获得更好解译精度。

1 数据获取与预处理

1.1 区域概况

本文选取石津灌区为试验区域,占地面积4 174.48 km²,其中耕地面积29万hm²,主要农作物为冬小麦、夏玉米及苹果、梨、桃等,为国家大II型灌区,位于河北省中南部,太行山东麓,滹沱河

下游以南、滏阳河以北、以西,覆盖石家庄、衡水、邢台3个市的14个县(市、区)。

1.2 数据选择

基于物候特征,冬小麦三月返青、四月拔节,此时影像中冬小麦与其他地物的光谱差异最为明显,因此选择三月底、四月初的高分影像初步识别冬小麦;夏玉米九月底全部收获,十月上旬开始播种冬小麦,夏玉米收获后其用地光谱特征在影像中表现为裸地,十月果树还未完全步入落叶期,依然表现为植被的光谱特征,故选择九月底或十月中上旬影像初步识别夏玉米和果树。为确保达到最佳解译结果,对比2014年3至4月、9至10月所有可获取高分数据的质量后,选择使用4月3日和10月16日两期高分影像。为保证时相匹配,从NASA网站下载2014年6月至2015年6月一整年MOD13Q1植被指数产品,共23幅影像,并对23幅影像进行编号,见表1。

表1 冬小麦、夏玉米、果树、建筑训练区NDVI时间序列特征值

Tab.1 NDVI time series characteristic values of winter wheat, summer maize, fruit trees and building training areas

日期	时相 编号	冬小麦训练区			果树训练区			夏玉米训练区			建筑训练区		
		最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值
2015.1.1	1	0.160 3	0.420 4	0.286 2	0.157 1	0.235 9	0.211 5	-	-	-	0.035 1	0.246 5	0.099 5
2015.1.17	2	0.151 8	0.364 3	0.259 6	0.184 7	0.242 2	0.223 3	-	-	-	0.034 7	0.261 3	0.106 8
2015.2.2	3	0.160 7	0.340 2	0.240 3	0.152 7	0.247 9	0.203 4	-	-	-	0.036 8	0.248 9	0.110 2
2015.2.18	4	0.160 0	0.342 9	0.244 0	0.163 4	0.238 9	0.217 7	-	-	-	0.052 4	0.214 2	0.115 4
2015.3.6	5	0.144 1	0.563 2	0.398 4	0.166 9	0.307 7	0.216 7	-	-	-	0.073 3	0.294 3	0.125 9
2015.3.22	6	0.211 6	0.518 0	0.394 7	0.187 0	0.320 6	0.215 3	-	-	-	0.087 8	0.394 6	0.143 2
2015.4.7	7	0.407 6	0.869 6	0.721 3	0.450 7	0.700 1	0.631 0	-	-	-	0.128 2	0.375 3	0.214 9
2015.4.23	8	0.501 4	0.875 7	0.773 2	0.565 8	0.795 1	0.747 2	-	-	-	0.163 4	0.643 7	0.300 7
2015.5.9	9	0.554 4	0.866 8	0.787 9	0.526 8	0.824 6	0.786 6	-	-	-	0.150 9	0.563 8	0.285 0
2015.5.25	10	0.259 4	0.653 7	0.374 4	0.526 4	0.801 8	0.715 8	-	-	-	0.138 4	0.504 1	0.263 9
2014.6.10	11	0.220 4	0.458 2	0.307 9	0.618 7	0.798 6	0.729 1	-	-	-	0.129 8	0.602 6	0.287 2
2014.6.26	12	0.237 3	0.555 9	0.313 6	0.509 1	0.817 5	0.760 8	0.226 9	0.614 4	0.286 7	0.154 3	0.556 0	0.281 5
2014.7.12	13	-	-	-	0.516 8	0.816 5	0.775 3	0.390 3	0.712 8	0.559 4	0.160 6	0.620 2	0.299 4
2014.7.28	14	-	-	-	0.613 3	0.845 0	0.808 5	0.622 9	0.947 3	0.835 9	0.052 4	0.729 8	0.337 8
2014.8.13	15	-	-	-	0.718 8	0.839 1	0.808 2	0.622 4	0.902 7	0.852 7	0.197 6	0.750 4	0.347 4
2014.8.29	16	-	-	-	0.638 1	0.851 8	0.819 8	0.505 0	0.901 7	0.828 4	0.158 0	0.770 8	0.321 0
2014.9.14	17	-	-	-	0.568 0	0.864 6	0.769 5	0.367 7	0.845 5	0.747 6	0.118 2	0.590 6	0.241 1
2014.9.30	18	-	-	-	0.604 4	0.806 2	0.756 0	0.182 5	0.762 5	0.355 9	0.061 9	0.701 3	0.311 5
2014.10.16	19	0.142 2	0.417 9	0.226 5	0.489 1	0.754 8	0.679 0	0.167 5	0.332 7	0.192 2	0.126 8	0.473 3	0.220 4
2014.11.1	20	0.187 6	0.474 6	0.341 8	0.328 1	0.475 1	0.435 5	-	-	-	0.085	0.416 8	0.184 2
2014.11.17	21	0.187 4	0.626 9	0.455 8	0.249 4	0.396 4	0.343 6	-	-	-	0.047 4	0.319 1	0.157 7
2014.12.3	22	0.188 9	0.571 7	0.390 6	0.196 5	0.322 7	0.292 6	-	-	-	0.151 7	0.323	0.128 0
2014.12.19	23	0.160 3	0.486 1	0.338 0	0.231 5	0.289 0	0.264 1	-	-	-	0.007 1	0.264 1	0.121 7

1.3 数据预处理

高分一号影像的数据预处理主要包括:辐射定标、大气校正、正射校正、影像拼接、研究区裁剪和几何校正。为使 MOD13Q1 数据与高分数据在地理位置上统一,下载 2014 年任意一期覆盖研究区的 Landsat8 影像,以 15 m 空间分辨率的 Landsat8 全色波段为基准,对高分一号影像进行几何校正。MODIS 影像数据预处理主要包括:投影变换、影像裁剪、数据重采样。MOD13Q1 产品 NDVI 的原始数据值范围是 $[-10\ 000, 10\ 000]$,需要将所有影像的 DN 值均乘以 0.000 1 变换为 NDVI 正常值范围 $[-1, 1]$ 。最后将变换值域后的 23 幅影像进行叠加,形成 MODIS NDVI 时间序列数据。

1.4 样本选择

通过目视解译方法,在 Google earth 中选取研究区范围内建筑、冬小麦、夏玉米及果树 4 类样本,并保证样本均匀分布和随机性,最终选取建筑样本 150 个,冬小麦样本 530 个,夏玉米样本 350 个,果树样本 130 个。其中,150 个建筑样本,400 个冬小麦样本,250 个夏玉米样本,80 个果树样本作为监督分类和构建时间序列的训练样本;130 个冬小麦样本,100 个夏玉米样本,50 个果树样本用来检验总体解译效果。

2 研究方法

首先使用最大似然法对预处理后的高分影像进行监督分类,以获取初步识别结果;然后根据训练样本统计绘制各作物 MODIS NDVI 变化曲线,分析 NDVI 的时间序列,选择区分作物的最佳时相并确定识别阈值;最后构建监督分类结果与 NDVI 时序共同约束下的决策树分类模型,识别出研究区主要作物的空间分布。

2.1 高分影像的监督分类

对 4 月 3 日和 10 月 16 日高分数据使用最大似然法进行监督分类,分类后结果中存在大量的细小图斑,需要对结果进行包括 Majority 分析、聚类处理及过滤处理在内的分类后处理。分类后处理可以去除监督分类初步结果中的细小图斑,并能够在一定程度上简化分类影像。

2.2 NDVI 时间序列的构建

使用目视解译出的各作物训练区裁剪 MODIS

NDVI 时间序列影像,结合 ENVI 软件统计工具,统计得到冬小麦、夏玉米、果树、建筑训练区的 NDVI 时间序列特征值,见表 1,由 NDVI 均值得出作物的生长趋势线,见图 1。

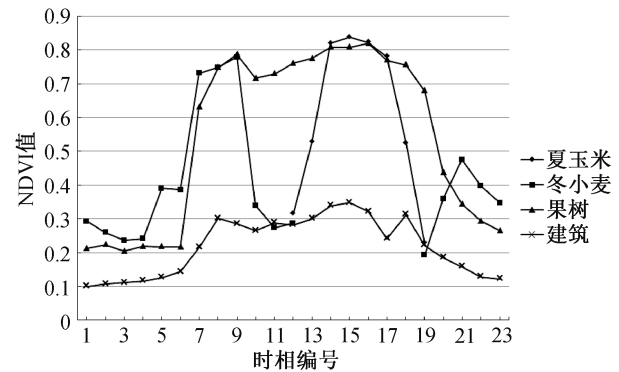


图 1 NDVI 变化曲线

Fig.1 NDVI changing curve

2.3 GF-1 监督分类结果与 MODIS NDVI 时间序列共同约束下的决策树构建

高分数据具有受气候因素影响,如云雨的覆盖影响较为严重的数据特性,导致监督分类中对作物的错分、漏分,从而造成总体解译结果的精度损失。以监督分类结果为约束的决策树构建能对错分、漏分像元进行筛选和剔除。影像分类阈值的确定对决策树的精准实施起决定性作用,本文的阈值指标对应 NDVI 的特征值。通过观察作物生长趋势线,找出区分不同作物的最佳识别点,使用不同时相 NDVI 最大或最小值作为阈值限制作物提取范围,争取达到某种作物的完整提取。

2.3.1 冬小麦决策树识别模型

由图 1 中冬小麦 MODIS NDVI 值随时间变化的曲线可以看出,序列号为 11 的时相是区分冬小麦的最佳时相,此时 NDVI 值处于主要作物中的最小值,针对该特征的约束条件为:第 11 时相的 NDVI 小于冬小麦第 11 时相的最大值。从序号为 9 的时相到第 12 时相,冬小麦 NDVI 值有一个明显的下降,是冬小麦生长周期中一个可与其他地物区分的明显特征,针对该特征的约束条件为:第 9 时相与第 10 时相 NDVI 之和大于第 11 时相与第 12 时相 NDVI 之和。由于 GF 影像较高的空间分辨率,观察 4 月 3 日的 GF-1 原始影像发现,此时是区分冬小麦与其他地物的最佳时相,可以保证监督分类冬小麦的最佳分类精度,故将监督分类结果作为首要约束,针对该特征的约束条件为:提

取4月3日高分影像监督分类等于冬小麦的部分。本文综合约束下的冬小麦决策树识别模型为:

$$GF(4.3) = \text{冬小麦}$$

$$NDVI(11) < 0.4582$$

$NDVI(9) + NDVI(10) > NDVI(11) + NDVI(12)$
其中GF(4.3)为4月3日高分影像监督分类结果NDVI(A)中A为第A时相。

2.3.2 果林决策树识别模型

结合监督分类结果观察原始高分影像发现,4月3日的监督分类结果中出现了果林和建筑的大面积错分,故在高分监督分类结果提取果树和建筑两类地物。根据NDVI变化曲线,第11时相的果树NDVI的最小值大于建筑NDVI的最大值,因此对于MODIS影像使用第11时相区分果树与建筑,针对该特征的约束条件为:第11时相的NDVI大于建筑NDVI的最大值。本文综合约束下的果树决策树识别模型为:

$$GF(4.3) = \text{建筑} + \text{果树}$$

$$NDVI(11) > 0.6178$$

2.3.3 夏玉米决策树识别模型

果树和夏玉米生长周期存在重叠期,所以首先在10月16日高分监督分类结果中将已确定的果树结果剔除。因为影像质量问题,在10月16日高分影像监督分类结果中,建筑和夏玉米出现严重的错分现象,故将夏玉米和建筑一同提取,再使用其他约束条件单独提取夏玉米。从MODIS NDVI时间序列曲线可以看出,夏玉米在第15时相NDVI达到最大,因为此时相正值夏季,是所有植物生长最旺盛的季节,使NDVI值大于夏玉米NDVI在第15时相的最小值,能提取出夏玉米及大部分同期生长植物并排除居民区;第12、13时相是夏玉米播种生长阶段,NDVI值较小,其他植物生长旺盛,NDVI值维持在较高水平,使用12、13时相的最大值共同约束,区分夏玉米于其他茂盛植物,针对该特征的约束条件为:第12、13时相的NDVI值小于该时相夏玉米NDVI最大值。高分数据和MODIS数据综合约束下的夏玉米识别模型为:

$$GF(10.16) \neq \text{果树}$$

$$GF(10.16) = \text{建筑} + \text{玉米}$$

$$NDVI(15) > 0.6224$$

$$NDVI(12) < 0.6144$$

$$NDVI(13) < 0.7128$$

其中GF(10.16)为10月16日高分影像监督分类结果。

3 结果与分析

3.1 GF影像监督分类结果与评价

基于GF-1监督分类的结果如图2所示。根据Google Earth选取的验证样本检验GF数据监督分类结果中的正确分类点,对4月3日和10月16日GF影像主要作物的分类结果进行精度评价,见表2。

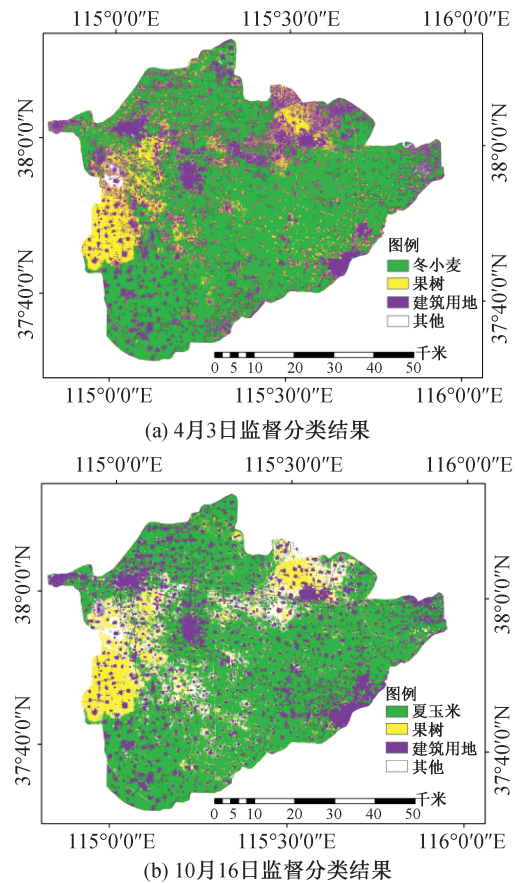


图2 监督分类结果

Fig.2 Supervised classification results

表2 GF-1 监督分类精度验证表

Tab.2 GF-1 accuracy verification table of supervised classification

日期	作物类型	验证样本个数	正确分类个数	分类精度/%
4月3日	冬小麦	130	119	91.58
	果林	50	43	86
10月16日	夏玉米	100	65	65
	果林	50	44	88

理想状况下,高空间分辨率影像与监督分类的结合可以获得较高的解译精度^[10],但通常受到所获取到的GF数据质量的影响,解译精度会有所

降低。从表 2 中可以看出,4 月 3 日影像上冬小麦和果林的分类精度相对较高,是因为该时相冬小麦和果林都有其独特的具有可分性的光谱特征;再结合 GF 原始影像,发现分类结果中存在部分将建设用地划分为果林的多分现象。从表中可以看出,10 月 16 日影像对夏玉米的解译效果较差,观察 GF 原始影像发现,气象条件对该时相作物光谱特征的影响较大,云雾、日照条件等使地物间可分性降低,最终导致夏玉米分类精度较其他地物的大幅度下降;又因该时相果林的光谱特征明显区别于其他地物,果林的分类精度仍保持在较高水平。

3.2 GF-1 监督分类与 NDVI 时间序列共同约束的决策树分类结果与评价

基于 GF-1 监督分类与 NDVI 时间序列共同约束的决策树产生的解译结果如图 3 所示。

3.2.1 基于训练验证点的分类精度验证

根据 Google Earth 选取的验证样本检验综合解译结果,分类精度见表 3。

表 3 基于 Google Earth 样本点的精度验证表
Tab.3 Accuracy verification table based on Google Earth sample points

作物类型	验证样本个数	正确分类个数	分类精度/%
冬小麦	130	124	95.38
夏玉米	100	91	91
果林	50	47	94

3.2.2 基于随机点的分类精度验证

使用 Arcgis 的随机点生成工具,在研究区内随机生成 200 个随机点,在综合解译方法的冬小麦、果林和夏玉米结果中分别进行统计,200 个点中共有 132 个点对应冬小麦,54 个对应果林,113 个对应夏玉米;再将这些点分别投映到 Google Earth 中,统计各类别点的正确分类个数,对综合方法的分类精度做出评价,分类精度见表 4。

表 4 基于随机点的精度验证表

Tab.4 Accuracy verification table based on random points

作物类型	对应个数	正确分类个数	分类精度/%
冬小麦	132	125	94.69
夏玉米	113	103	91.15
果林	54	50	92.59

由表 3 和表 4 可见,两种验证方法的结果基本一致,都能观察到分类精度的显著提高。基于监督分类与决策树的综合分类方法的总解译精度达

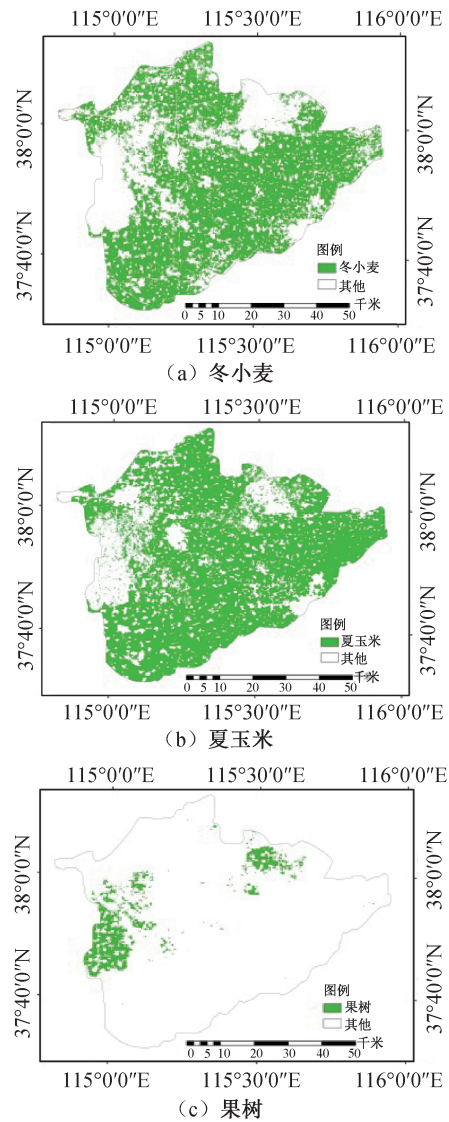


图 3 多种约束决策树作物解译结果
Fig.3 Crop interpretation results of multiple constraint decision trees

93%以上,相比较于单纯的监督分类,冬小麦分类解译提高 3%以上,果树解译精度提高 8%以上,夏玉米解译精度提高 26%,总体精度提高了 10%以上。

4 结论

1) MODIS NDVI 时间序列和 GF 影像监督分类结合的决策树综合分类方法的精度较高,总体分类精度达到 93.13%,较 4 月 3 日的 GF 影像监督分类,精度提高了 4%以上,较 10 月 16 日的 GF 影像监督分类,精度提高了 10%以上,说明该方法具有较高的识别精度,能更好地反映区域主要作物的种植分布,适合作物的分类和提取。

2)本研究综合方法的夏玉米的解译精度较单纯的监督分类有了很大的提升,因为10月16日的GF影像质量不利于区分夏玉米,导致直接监督分类夏玉米的可行性降低,只关注夏玉米识别精度的提升,并不能体现出该方法在解译精度方面的优越性。高分影像监督分类的冬小麦和果林的解译已经取得较好效果,在使用综合解译方法后精度依旧有4%以上的提高,证明该方法可一定程度克服GF数据质量的影响,并对GF影像的监督分类结果有一定的改善、修正效果。

3)GF影像和MODIS数据的综合使用,使数据更好地发挥各自的优势并在时空分辨率上相互补充,提高了数据的可用性和应用潜力。

本文在选取样本时,采用了目视解译的方法,这会对遥感解译结果造成一些人为的误差。由于本文研究区域跨多个市县,尚缺少对该区域的完整的作物种植信息的统计数据,故精度上未采用统计数据验证,对于其他区域,可增加统计数据验证分类精度,对本方法进行更加全面的精度评价。

参考文献:

- [1]唐华俊,吴文斌,杨鹏,等.农作物空间格局遥感监测研究进展[J].中国农业科学,2010,43(14):2879-2888.
- [2]张莉,吴文斌,杨鹏,等.黑龙江省宾县农作物格局时空变化特征分析[J].中国农业科学,2013,46(15):3227-3237.
- [3]BASTIAANSEN W G M. Remote Sensing in Water Resources Management: the State of the Art[M]. Colombo, Sri Lanka: IWMI Press, 1998: 118.
- [4]胡琼,吴文斌,宋茜,等.农作物种植结构遥感提取研究进展[J].中国农业科学,2015(10):34-48.
- [5]曹卫彬,杨邦杰,宋金鹏. TM影像中基于光谱特征的棉花识别模型[J].农业工程学报,2004,20(4):112-116.
- [6]YANG C, EVERITT J H, MURDEN D. Evaluating High Resolution SPOT 5 Satellite Imagery for Crop Identification[J]. Computers and Electronics in Agriculture, 2011, 75(2):347-354.
- [7]BECK P S A, ATZBERGER C, HOGDA K A, et al. Improved Monitoring of Vegetation Dynamics at very High Latitudes: A New Method Using MODIS NDVI[J]. Remote Sensing of Environment, 2006, 100(3):321-334.
- [8]杨小唤,张香平,江东.基于MODIS时序NDVI特征值提取多作物播种面积的方法[J].资源科学,2004,26(6):17-22.
- [9]许青云,杨贵军,龙慧灵,等.基于MODIS NDVI多年时序数据的农作物种植识别[J].农业工程学报,2014,30(11):134-144.
- [10]刘克宝,刘述彬,陆忠军,等.利用高空间分辨率遥感数据的农作物种植结构提取[J].中国农业资源与区划,2014,35(1):21-26.

(责任编辑 王利君)