黄晓君，颉耀文，卫娇娇，等. 基于变化检测 - CART 决策树模式自动识别沙漠化信息[J]. 灾害学，2017，32(1)：36 - 42.［HUANG Xiaojun，XIE Yaowen，WEI Jiaojiao，et al. Automatic recognition of desertification information based on the pattern of change detection-CART decision tree［J］. Journal of Catastrophology，2017，32(1)：36 - 42. doi：10.3969/j.issn.1000 - 811X.2017.01.008.］

基于变化检测 - CART 决策树模式自动识别沙漠化信息*

黄晓君[1,2]，颉耀文[1]，卫娇娇[1]，付 苗[2]，吕利利[1]，张玲玲[1]

（1. 兰州大学 资源环境学院，甘肃 兰州 730000；
2. 内蒙古师范大学 地理科学学院，内蒙古 呼和浩特 010022）

摘 要：目前沙漠化遥感监测存在目视解译的局限性、数据源的约束性、遥感信息利用率低等问题。基于此，以民勤盆地为试验区，首先采用图像差值、最大值合成及二维最大类间方差等方法，检测 1994 年、2014 年两期 Landsat 图像的变化像元，然后利用分类与回归树(CART)算法构建决策树，自动提取了 2014 年沙地信息，最后将变化检测结果与沙地信息进行空间叠置分析，并实现了沙漠化信息自动识别模式。研究表明，变化检测 - CART 决策树模式精度为 89.43% ~93.00%，在 95% 置信水平上其置信区间介于 85.90% ~98.00%，显然其精度具有较高可信度；该模式不仅能够充分利用丰富遥感信息而且可排除多余信息的干扰。可见，变化检测 - CART 决策树模式是识别沙漠化信息的有效方法之一，将对沙漠化防治工程具有重要应用价值。

关键词：沙漠化；分类与回归树(CART)；决策树；变化检测；自动识别

中图分类号：X43；F301 **文献标志码**：A **文章编号**：1000 - 811X(2017)01 - 0036 - 07

doi：10.3969/j.issn.1000 - 811X.2017.01.008

土地沙漠化是指在干旱、半干旱和部分半湿润地区，由于自然因素或人文因素或两者的共同作用，使原非沙漠地区出现以风沙活动为主要标志的类似沙漠景观的土地退化过程。它是当今世界最为严重的环境 - 社会 - 经济问题之一，及时准确监测沙漠化信息是治沙止漠的基础，研究意义重大。

遥感具有获取信息快、监测范围大、提供信息丰富等特点，1970 年代以来在土地沙漠化中得到广泛应用。国外最早在 1975 年 Lampery 利用航空像片对撒哈拉沙漠南缘的移动进行了研究[1]，而国内 1984 年中国科学院沙漠研究所利用航空相片及卫星像片研究沙漠化[2]。目前，常见的土地沙漠化遥感监测方法主要有目视解译[3-5]、遥感专题指数[6-8]、沙漠化监测指数[9-10]、人工神经网络[11-12]、光谱混合分析[13-14]、决策树分类[15-16]、基于面向对象[17-18]等。由于沙漠化过程的动态性与成因的复杂性，快速准确监测沙漠化信息有一定挑战性。目视解译提取沙漠化虽精度较高，但工作量大、效率低、主观性强，很难及时掌握沙漠化状况。遥感专题指数、沙漠化监测指数、人工神经网络等方法对遥感信息的利用率较低，其丰富的信息未能充分发挥应有的作用。光谱混合分析与基于面向对象的监测虽精度较高，但对数据源要求很高，如光谱混合像元分析较适于高光谱遥感数据，基于面向对象适于高空间分辨率遥感数据，从而它们利用的遥感数据不仅价格昂贵而且无法满足大面积多时相的需求。目前虽沙漠化遥感监测方法较多，但普适性较差，遥感信息的利用率低，尚未建立普遍接受的自动提取沙漠化信息的方法框架。

分类与回归树(CART)算法不仅对遥感数据源要求不高，而且能够充分利用其丰富的遥感信息。目前美国 NASA 陆地卫星 Landsat 可供免费的 1980 年代以来大面积、多时相、中高空间分辨率多光谱遥感数据，它必然是沙漠化遥感监测的主要数据源。基于此，本研究以民勤盆地为试验区，利用 Landsat 卫星遥感数据，尝试变化检测 - CART 决策树模式是构建计算机自动识别沙漠化信息的有效方法。

* 收稿日期：2016 - 05 - 17 修回日期：2016 - 07 - 04
基金项目：国家自然科学基金项目(41471163)；兰州大学中央高校基本科研业务费专项资金(lzujbky - 2016 - 242)
第一作者简介：黄晓君(1984 -)，男，蒙古族，内蒙古兴安盟人，讲师，主要从事自然灾害遥感监测与预警研究.
E-mail：hxj3s@qq.com

1　试验区及材料

本文试验区是甘肃民勤盆地，地处我国西北内陆干旱区，在河西走廊东北部石羊河下游，地理位置介于 38°25′～39°17′N，102°45′～104°15′E。盆地底部海拔在 1 300～1 400 m 之间，大部平坦开阔，地势由西南向东北缓倾。盆地东西北三面被腾格里和巴丹吉林两大沙漠包围，大陆性沙漠气候特征明显，干旱少雨、昼夜温差大、光照充足。在盆地石羊河两岸分布所谓民勤绿洲。多年来，由于人类活动的加剧和气候变化的影响，绿洲与沙漠交错带出现了土地沙漠化现象，因此该地区可作为本研究试验区。

本文主要使用的数据有 1994 年与 2014 年 7 月份 Landsat TM/OLI 影像数据（云量小于 2%，行列号为 131033，空间分辨率为 30 m），还有美国 NASA SRTM1 DEM 数据（空间分辨率为 30 m），另外包括 2014 年 7 月与 10 月两次野外调查数据。利用 ENVI 软件对影像数据进行大气辐射校正和裁剪处理，对 DEM 数据进行几何校正、投影转换及裁剪处理。影像、DEM 数据的坐标与投影系统应统一，即坐标系统为 WGS1984，投影为通用墨卡托 UTM，带号为 48 N。

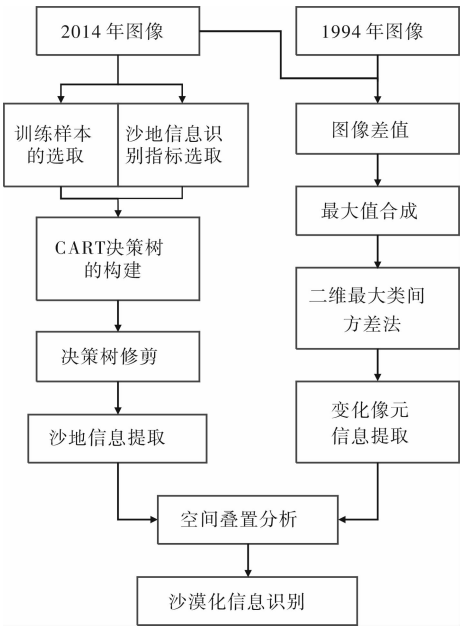


图 1　总体技术路线图

2　研究方法

基于变化检测－CART 决策树模式自动识别沙漠化信息的总体思路如图 1 所示，首先自动检测由 1994 年至 2014 年的土地变化像元，它是图像差值结合最大值合成法及二维最大类间方差法来完成；其次自动提取 2014 年沙地信息，它是在获取若干沙地信息识别指标的基础上，利用 CART 算法构建决策树来完成；最后自动识别沙漠化信息，它是通过变化检测结果与沙地信息的空间叠置分析来实现。

2.1　变化检测

2.1.1　图像差值

为分析 1994 年、2014 年 TM/OLI 图像蓝、绿、红、近红外及短波红外 1、2 波段对典型地类，如沙地、绿地（如草地、耕地、林地）、水域（如河渠、湖泊、水库）及其他土地（如居民地、工矿、道路、盐碱地、裸地、戈壁）等的光谱反射曲线特征，采样了以上地类的反射率，并绘制了光谱反射曲线（图 2）。结果表明，沙地的反射率在红、短波红外 1、2 波段上突出，与周围的绿地、水域及其他土地的光谱反射曲线有明显差异，因为红色波段对叶绿素有较强的吸收，而对沙地的反射较高，且它是可见光中信息量较丰富的波段；短波红外 1、2 波段对水体有较强的吸收，而对沙地的反射高，且在所有波段中短波红外 1 的信息量最丰富。因此，从两期图像中分别选择了近红外、短波红外 1、2 等 3 个波段，进行了相应的差值处理。

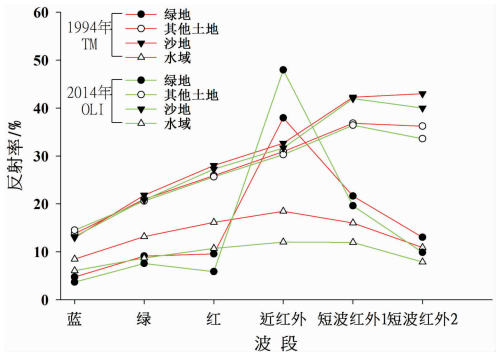


图 2　典型地类光谱反射曲线

2.1.2　变化像元信息提取

图像差值为分离变化与未变化像元信息奠定了基础，因为差值后像元绝对值呈两极分化态势，这是变化像元信息提取的切入点。本研究顾及压缩数据和不遗漏变化信息，利用最大值合成法，将 3 个差值图像进行合成处理。另外，为减少多余信息对变化像元中沙漠化信息的干扰，从合成图像中剔除小于等于 0 的像元，因为图像差值使沙漠化像元值基本控制在大于 0 范围。在此基础上，采用二维最大类间方差法，提取合成图像中的变化像元信息。二维最大类间方差法是在二维灰度直方图的基础上，具有统计意义上的最佳分割阈值，其将图像的二维灰度直方图分割成两类，使两类间离散度矩阵的迹取最大值，即分离性最大，如式（1）所示。

$$t_r\rho(\alpha',\beta')=\max_{1<\alpha,\beta<L}\{t_r\rho(\alpha,\beta)\}。\qquad(1)$$

表 1 沙漠化信息识别指标的计算

沙地信息识别指标	表达式	参考来源
MSAVI	$\rho_{nir}+0.5=\sqrt{(\rho_{nir}+0.5)^2-2(\rho_{nir}-\rho_{red})}$	Qi J[19]
NDVI	$(\rho_{nir}-\rho_{red})/(\rho_{nir}+\rho_{red})$	Tucker C J[20]
NDWI	$(\rho_{green}-\rho_{swir1})/(\rho_{green}+\rho_{swir1})$	Mcfeeters S K[21]
NDBI	$(\rho_{swir1}-\rho_{nir})/(\rho_{swir1}+\rho_{nir})$	吴宏安[22]
SI	$\sqrt{\rho_{blue}\times\rho_{red}}$	Khan N M[23]
GSI	$(\rho_{red}-\rho_{blue})/(\rho_{red}+\rho_{green}+\rho_{blue})$	Xiao J[24]
Albedo	$0.356\rho_{blue}+0.130\rho_{red}+0.373\rho_{nir}+0.085\rho_{swir1}+0.072\rho_{swir2}-0.0018$	Liang Shunlin[25]
Brightness Greenness	K－T 变换(缨帽变换)	－
纹理特征	二阶概率统计的滤波	－

备注：ρ_{blue}、ρ_{green}、ρ_{red}、ρ_{swir1}、ρ_{swir2}分别为蓝、绿、红、近红外及短波红外 1、2 等波段的反射率

式中：$t_r\rho$ 为类间离散度矩阵 ρ 的迹，L 为图像灰度级，(α, β)为任意阈值，(α', β')为最佳阈值。

本文根据以上数学定义，实现了变化像元与未变化像元的有效分离，其结果显示合成图像像元值大于 0.076 为变化像元，反之则未为变化像元。

2.2 CART 决策树沙地信息提取

2.2.1 沙地信息识别指标选取

目前尚没有能够直接准确提取沙地信息指数的方法，一般运用分类方法提取。本研究首先把地表分为沙地、绿地(如草地、耕地、林地)、水域(如河渠、湖泊、水库)及其他土地(如居民地、工矿、道路、盐碱地、裸地、戈壁)等 4 类，然后利用 CART 决策树对地表进行分类，并实现沙地信息提取。

笔者参考文献[7、9、15]等，选择修改型土壤调整植被指数(MSAVI)、归一化植被指数(NDVI)、归一化差异水体指数(NDWI)、归一化裸露指数(NDBI)、盐分指数(SI)、表土粒度指数(GSI)、地表反照率(Albedo)、K－T 变换的分量(如亮度 Brightness、绿度 Greenness)、纹理特征及 DEM 等作为沙地信息识别指标。上述指标在分类中发挥应有的作用，如：MSAVI、NDVI、Greenness 能够表征地面绿地信息，其中 NDVI 由于对低密度植被覆盖地区受土壤背景影响较大，且高密度植被覆盖地区易饱和，故适合中密度植被覆盖地区绿地信息识别，而 MSAVI 受土壤背景影响较小，适于低密度植被覆盖地区的绿地信息识别，Greenness 与绿色植物的数量具有密切关系，能够弥补 NDVI 在高密度植被覆盖地区快速饱和的缺陷；NDWI 可以最大程度抑制植被信息，而能够突出水体信息；Albedo 能够反映地表对太阳辐射反射特性，随地表植被覆盖度、土壤水分及地表粗糙度的变化，Albedo 具有明显的响应，则沙地、绿地、水域等不同地表特征的反射率呈明显差异；Brightness 能够反映土壤反射率变化的信息，受土壤水分的影响，其值对不同水分地表具有明显差异，则分离沙地与其他土地起到应有的作用；NDBI 对裸露地表较敏感，则识别裸地、道路等信息较好；SI 能够反映盐碱地信息；GSI 对表土粒度信息较敏感，则能够较好识别戈壁、居民地等信息；纹理特征能够反映地表粗糙度信息，沙地纹理与其他地表具有明显的差异，则识别沙地信息起到应有的作用；DEM 能够反映地貌形态特征，沙地与戈壁在海拔上具有一定空间差异性，则其能够增强沙地与戈壁的分离性。以上指标的获取如表 1 所示。

2.2.2 训练样本的选取

本文将地表分为 4 种土地类型，即沙地、绿地(如草地、耕地、林地)、水域(如河渠、湖泊、水库)及其他土地(如居民地、工矿、道路、盐碱地、裸地、戈壁)，并采用目视解译与野外考察等手段，在 2014 年图像上选取训练样本(350 个图斑，共 28 138 个像元)。目视解译是在解译者具备图像解译专业知识、区域背景知识及遥感系统知识的基础上，根据图像基本解译要素和解译标志来识别目标地物。训练样本的选取是构建决策树的关键部分，它的好坏直接影响后期分类精度。因此，在选取训练样本时，考虑了其代表性、典型性及完备性，也就是说训练样本代表了整个区域内各地类光谱特征差异。

2.2.3 CART 决策树的构建及分类

本文利用 CART 算法建立了决策树，即称为 CART 决策树。CART 决策树是由输入变量和输出变量构成的训练样本数据集的循环分析的二叉树结构，其中输入变量是由基于训练样本像元的各沙地信息识别指标产生；输出变量是由基于训练样本像元的 4 种土地类型产生。CART 算法包括决策树生长和决策树修剪两部分内容。在决策树生长时，利用杂度削减量从众多输入变量中选择当前最佳分组变量，然后采用基尼系数确定分组变量分割阈值，其数学定义如式(2)、式(3)所示。

$$G(t)=1-\sum_{j=1}^{k}P^2(j\mid t)，（其中\ P(j\mid t)=\left(\frac{N_{j,t}}{N_j}\right)/\sum_j\frac{N_{j,t}}{N_j}）。\quad(2)$$

式中：$G(t)$为基尼系数，t 为节点，k 为输出变量类别个数，$P(j\mid t)$是节点中输出变量取第 j 类的归一化概率，$N_{j,t}$是节点 t 包含第 j 类的样本量，N_j是全样本量。分组变量分割阈值过程是首先将每个输入变量数值按升序排序，再从小到大依次以相邻数值的中间值作为分割点，将样本分成两组，并计算两组样本输出变量值的异质性。当使两组输出变量值的异质性总和最小、纯度最大时，分组变量达到最佳分割点。输出变量值的异质性是用基尼系数来测度，当节点中输出变量均取同一类别时，输出变量取值的异质性最小，即基尼系数为 0，而当各类别取值概率相等时，输出变量取值的异质性最大，即基尼系数为 $1-1/k$。

$$\Delta G(t)=G(t)-\frac{N_r}{N}G(t_r)-\frac{N_l}{N}G(t_l)。\quad(3)$$

式中：$\Delta G(t)$为杂度削减量，$G(t)$和 N 分别为分组前输出变量的基尼系数和样本量，$G(t_r)$、N_r和 $G(t_l)$、N_l分别为分组后右子树的基尼系数和样本量以及左子树的基尼系数和样本量。在节点处首先对所有输入变量将计算其杂度削减量，再对杂度削减量进行比较，选择最大对应的输入变量将成为当前节点分组变量。

在 CART 决策树构建中，决策树的修剪处理很重要。它在保证决策树有一定预测精度的前提下，还能削减决策树的复杂程度，最终获得一颗理想的树。修剪由预修剪和后修剪两部分构成，其中预修剪是在建立决策树时，事先制定一些控制参数来修剪决策树，如决策树最大深度、树中父节点和子节点的最少样本量或比例、树节点中输出变量的最小杂度削减量等；后修剪是在决策树预修剪后充分生长的基础上，根据一定的规则，修剪树中那些不具有一般代表性的叶节点或子树，这是一个边修剪边预测当前决策子树精度的过程。它是采用最小代价复杂性修剪法来完成。此方法不仅顾及了决策树的预测精度，而且考虑了决策树的复杂程度。本文通过上述修剪处理，得到理想的决策树，再利用该决策树自动提取了 2014 年沙地信息。

2.3　沙漠化信息识别

通过 1994 年、2014 年两期 Landsat 卫星图像变化检测与 2014 年 CART 决策树分类结果的空间叠置分析，自动识别了 2014 年沙漠化信息。若变化像元与分类图像的沙地像元相重合，当前像元为沙漠化像元，反之则未沙漠化像元。利用 ENVI 软件波段计算功能进行空间叠置分析，实现了沙漠化信息的自动识别（如图 3 所示，2014 年民勤盆地沙漠化信息识别的局部放大图）。

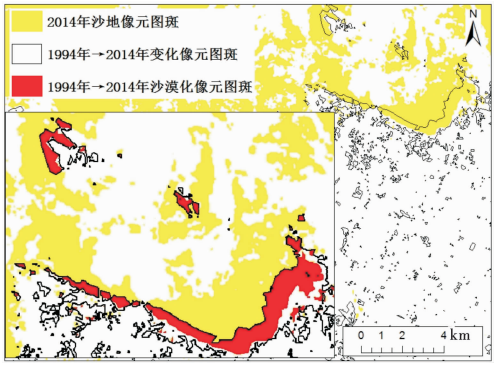


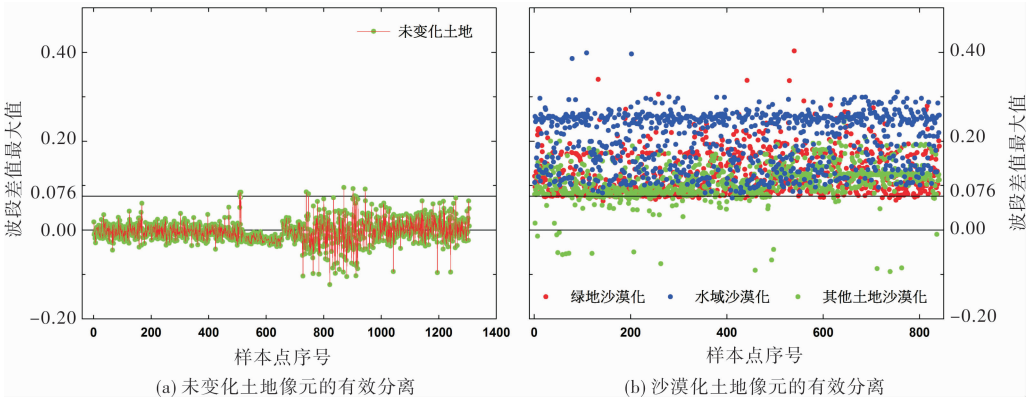
图 3　沙漠化信息识别的局部放大图

3　研究结果与分析

3.1　变化检测的有效性分析

3.1.1　沙漠化土地与未变化土地像元的有效分离

为分析变化检测对沙漠化与未变化土地像元能够分离的有效性，选取 1 300 未变化土地样本像元（覆盖了沙地、绿地、水域、其他土地）、2 520 沙漠化土地样本像元（绿地沙漠化、水域沙漠化及其他土地沙漠化像元各 840 个），然后分别对两期红、短波红外 1、2 波段进行差值，再进行最大值合成处理，获得其最大值，如图 4 所示。图 4a 和



(a) 未变化土地像元的有效分离　　(b) 沙漠化土地像元的有效分离

图 4　沙漠化与未变化土地像元的有效分离

4b 分别为未变化土地和沙漠化土地的 3 波段求差的最大值散点图，其表明：未变化土地像元值集中于 0 附近，其值基本都小于分割阈值(0.076)，而沙漠化土地像元值基本都大于分割阈值，从而在变化检测中最大值合成法和二维最大类间方差法的应用不仅克服变化信息丢失而且有效分离沙漠化与未变化像元。另外，试验显示有时其他土地沙漠化像元值小于阈值的现象，这对变化检测带来一定的误差，但这只是很小部分，可忽略不计。

3.1.2 变化检测的精度分析

为变化检测结果的精度分析，在 2014 年图像上随机选取了 3 154 个样本点，再利用目视解译与野外判别方法确定其土地类型，即沙地、绿地、水域及其他土地，并作为 2014 年真实样本点。在此基础上，同样的方法产生 1994 年相应真实样本点。对两期真实样本点进行比较得知，沙漠化与未变化样本点分别为 1 849 和 1 305，然后采用变化检测方法对两期真实样本点间进行检测，结果显示：沙漠化与未变化样本点分别为 1 613 和 1 286。显然沙漠化和未变化像元的检测精度分别为 87.24% 和 98.54%，总体精度为 91.92%(图 5)。

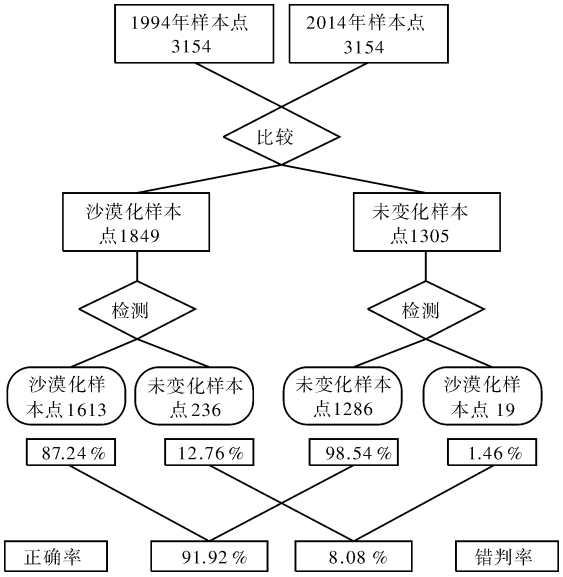


图 5 变化检测精度验证

3.2 CART 决策树沙地信息提取精度分析

3.2.1 CART 决策树模型风险评价

CART 决策树沙地信息提取精度分析之前，应评价 CART 决策树模型的精度，即模型风险评价。它是利用风险估计(错判率)来衡量。在 CART 决策树的构建中使用的总样本为 28 138 个像元。其中训练样本和测试样本分别占总样本的 80%、20%。一般用测试样本来测试模型精度。其测试结果显示，决策树对测试样本的风险估计为 0.031，即错判率为 3.1%，与训练样本的风险估计值相等(表 2)。说明该模型稳定，精度较高。

3.2.2 CART 决策树沙地信息提取精度评定

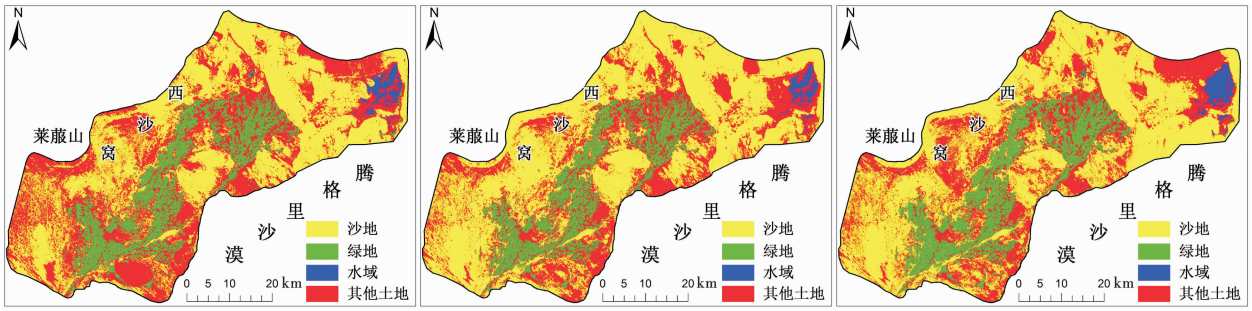
为评定 CART 决策树沙地信息提取精度，其分类结果与传统的最大似然法与现代的人工神经网络方法分类结果分别进行了精度比较，三者分类结果如图 6 所示。笔者利用混淆矩阵计算 Kappa 系数和总体精度来评定其分类精度。首先在 2014 年图像上随机选取了 260 个验证图斑；其次采用手持 GPS 导航找到验证图斑的实地位置，并确认该图斑土地类型；最后把确认的 260 个验证图斑作为真实感兴趣区，与相应的 CART 决策树、最大似然法、人工神经网络分类结果分别相比较，并获得真实与预测的混淆矩阵，同时求出 Kappa 系数和总体精度。结果显示，CART 决策树的分类总体精度和 Kappa 系数最高，分别为 87.34% 和 0.827 2，人工神经网络次之，分别为 85.71% 和 0.806 0，最大似然法最低，分别为 82.27% 和 0.758 8。分类总体精度是分类器正确分类的像元总数与总像元数之比，若总体精度越大，分类器分类精度越高；Kappa 系数是综合了生产精度和用户精度的指标，若 Kappa 系数越大，分类结果与真实参考之间差异越小，则分类器分类精度越高。显然 CART 决策树提取沙地信息优于最大似然法和人工神经网络。

3.3 沙漠化信息识别的精度验证

为沙漠化信息自动识别精度的验证，从变化检测 - CART 决策树模式的自动识别结果中随机选取了 1 486 个验证点(图 7d)，采用野外调查和室内目视解译方法，再结合地理国情普查数据和高空间分辨影像图，判别当前验证点是否沙漠化，如图 7a、图 b、图 c 显示了变化检测 - CART 决策树模式准确识别沙漠化的状况。精度验证得知，

表 2 模型风险的评价结果

预测	训练样本集				测试样本集			
实际	其他土地	水域	沙地	绿地	其他土地	水域	沙地	绿地
其他土地	5895	0	281	42	1505	0	72	4
水域	23	5487	0	0	4	1382	0	1
沙地	336	0	7369	0	91	0	1877	0
绿地	5	0	0	3054	1	0	0	709
总计	6259	5487	7650	3096	1601	1382	1949	714
风险估计	0.031				0.031			



(a) CART　　(b) 最大似然法　　(c) 人工神经网络

图 6　分类结果

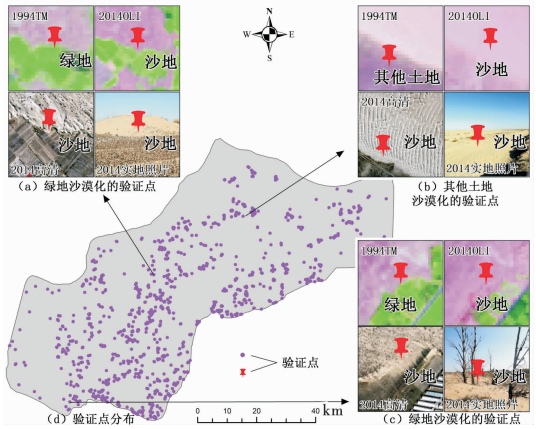


图 7　沙漠化验证

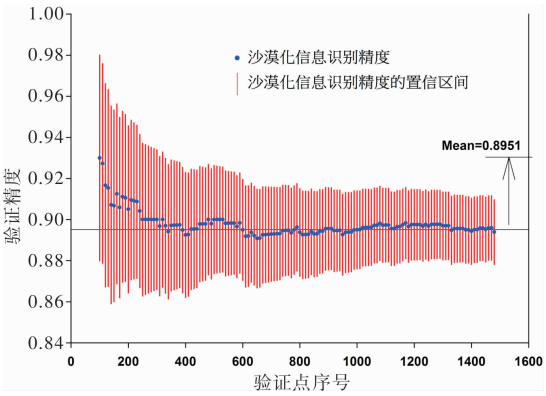


图 8　沙漠化信息识别精度及其置信区间

变化检测－CART 决策树精度为 89.43%（正确识别验证点数占总验证点数的百分比）。为进一步评定沙漠化信息识别精度的可信程度，在计算识别精度的基础上，利用总体比例的区间估计方法，在 95% 置信水平上估计了当前精度的置信区间（图 8），结果表明：随验证点由 100 增加到 1 486 期间，该模式沙漠化信息识别精度为 89.43% ~ 93.00%，其置信区间介于 85.90% ~98.00%；当验证点数增加到 600 之后，精度变化变为稳定，基本在 89.51% 上下微小浮动。综上可见，变化检测－CART 决策树模式自动识别沙漠化信息精度具有较高的可靠性。

4　结束语

本文以民勤盆地为试验区，提出了变化检测－CART 决策树模式自动识别沙漠化信息的方法。它由变化检测与 CART 决策分类树两部分组成，其中变化检测的主要作用是自动检测包含沙漠化信息的变化像元，其精度分析显示：沙漠化和未变化像元的检测精度分别为 87.24% 和 98.54%，总体精度为 91.92%，可见变化检测能够有效分离沙漠化与未变化像元。CART 决策树分类目的是从 2014 年图像中自动提取沙地信息，其精度分析表明：模型精度为 96.90%，分类总体精度为 87.34%，Kappa 系数为 0.827 2，与最大似然法、人工神经网络等方法比较得知，CART 决策树分类精度优于它们，显然 CART 决策树能够有效提取沙地信息。在变化检测与 CART 决策树分类的基础上，利用 ENVI 波段计算功能对变化检测与 CART 决策树分类结果进行空间叠置分析，并自动识别了沙漠化信息，其精度验证显示：变化检测与 CART 决策树模式精度为 89.43% ~93.00%，其置信区间介于 85.90% ~98.00%。显而易见，变化检测－CART 决策树模式是识别沙漠化信息的有效方法之一。

变化检测－CART 决策树模式不仅在精度上具有较高的可信度，而且其他方面亦有突出的优势，如：①与光谱混合分析、基于面向对象方法比较，适用性相对较高，它对数据源要求不高，免费 Landsat 卫星多光谱遥感数据即可，适合大面积沙漠化监测，但光谱混合分析和基于面向对象方法对遥感数据的光谱分辨率或空间分辨率具有较高的要求，则大面积沙漠化监测必然费用昂贵；②使丰富遥感信息充分发挥了应有的作用，而支持向量机、人工神经网络、遥感专题指数、沙漠化监测指数等方法对遥感信息的利用率相对较低；③实现了计算机自动识别沙漠化信息，排除了多余信息的干扰，避免了由于解译人员的主观性或经验的不同而引起的差异和错误，提高了工作效

率，可及时掌握沙漠化状况。可见，变化检测 - CART 决策树模式将对沙漠化防治工程具有重要应用价值。

参考文献：

[1] Lamprety H F. Report on the desert encroachment reconnaissance in northen Sudan[R]. Khartoum, Sudan National Council for Research, Ministry of Agriculture, 1975.

[2] 朱震达. 中国沙漠化研究的进展[J]. 中国沙漠, 1989, 9(1): 1-13.

[3] Reiche M, Funk R, Zhang Z D, et al. Application of satellite remote sensing for mapping wind erosion risk and dust emission-deposition in Inner Mongolia grassland, China[J]. Grassland Science, 2012, 58(1): 8-19.

[4] 王涛, 宋翔, 颜长珍, 等. 近35年来中国北方土地沙漠化趋势的遥感分析[J]. 中国沙漠, 2011, 31(6): 1351-1356.

[5] 胡梦珺, 田丽慧, 张登山, 等. 遥感与GIS支持下近30年来青海湖环湖区土地沙漠化动态变化研究[J]. 中国沙漠, 2012, 32(4): 901-909.

[6] Nasem Badreldin · Rudi Goossens. A satellite-based disturbance index algorithm for monitoring mitigation strategies effects on desertification change in an arid environment[J]. Mitig Adapt Strateg Glob Change, 2015, 20: 263-276.

[7] 孙建国, 姜烨, 颜长珍, 等. 利用专题指数改善沙漠化土地遥感分类精度[J]. 遥感技术与应用, 2013, 28(4): 655-658.

[8] Symeonakis E, Drake N. Monitoring desertification and land degradation over sub-saharan africa[J]. International Journal of Remote Sensing, 2004, 25(3): 573-592.

[9] 李宝林, 周成虎. 东北平原西部沙地沙质荒漠化的遥感监测研究[J]. 遥感学报, 2002, 6(2): 117-122.

[10] 曾永年, 向南平, 冯兆东, 等. Albedo-NDVI特征空间及沙漠化遥感监测指数研究[J]. 地理科学, 2006, 26(1): 75-77.

[11] 乔平林, 张继贤, 林宗坚. 基于神经网络的土地荒漠化信息提取方法研究[J]. 测绘学报, 2004, 33(1): 58-62.

[12] 买买提·沙吾提, 塔西甫拉提·特依拜, 丁建丽, 等. BP神经网络的沙漠化土地信息提取研究[J]. 干旱区研究, 2008, 25(5): 647-652.

[13] Dawellbait M, Morari F. Monitoring desertification in a savannah region in sudan using Landsat images and spectral mixture analysis[J]. Journal of Arid Environments, 2012, 80: 45-55.

[14] 李晓松, 吴波, 范文义, 等. 基于光谱混合分析的荒漠化信息提取——以毛乌素沙地为例[J]. 林业科学研究, 2006, 19(2): 192-198.

[15] 贾树海, 韩志根, 吕默楠, 等. 基于决策树的辽宁省北部沙漠化信息提取研究[J]. 生态环境学报, 2011, 20(1): 13-18.

[16] 段英杰, 何政伟, 王永前, 等. 基于遥感数据的西藏自治区土地沙漠化监测分析研究[J]. 干旱区资源与环境, 2014, 28(1): 55-61.

[17] 王志波, 高志海, 王琦瑜, 等. 基于面向对象方法的沙化土地遥感信息提取技术研究[J]. 遥感技术与应用, 2012, 27(5): 770-777.

[18] 冯益明, 郑冬梅, 智长贵, 等. 面向对象的沙化土地信息提取[J]. 林业科学, 2013, 49(1): 126-133.

[19] Qi J, Chehbouni A, Huete R, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment. 1994, 48(2): 119-126.

[20] Tucker C J, Justice C O, Prince S D. Monitoring the grasslands of the sahel 1984-1985[J]. International Journal of Remote Sensing, 1986, 7(11): 1571-1581.

[21] Mcfeeters S K. The use of normalized difference water index(NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.

[22] 吴宏安, 蒋建军, 周杰, 等. 西安城市扩张及其驱动力分析[J]. 地理学报, 2005, 60(1): 143-150.

[23] Khan N M, Rastoskuev V V. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators[J]. Agricultural Water Management, 2005, 77(3): 96-109.

[24] Xiao J, Shen Y, Tateishi R, et al. Development of topsoil grain size index for monitoring desertification in arid land using remote sensing[J]. International Journal of Remote Sensing, 2006, 27(12): 2411-2422.

[25] Liang Shunlin. Narrowband to broadband conversion of land sueface albedo I algorithms[J]. Remote Sensing of Environment, 2001, 76(2): 213-238.

Automatic Recognition of Desertification Information Based on the Pattern of Change Detection-CART Decision Tree

HUANG Xiaojun[1,2], XIE Yaowen[1], WEI Jiaojiao[1], FU Miao[2], LYU Lili[1] and ZHANG Lingling[1]

(1. *College of Earth and Environmental Sciences, Lanzhou University, Lanzhou* 730000, *China*;
2. *College of Geographical Science, Inner Mongolia Normal University, Huhhot* 010022, *China*)

Abstract: The desertification of remote sensing monitoring has some problems, such as visual interpretation limitation, constraint of data source and low utilization rate of remote sensing information. Based on this, taking Minqin basin as the test area, firstly, this paper detects the change pixel of two Landsat images in 1994 and 2014, by employing the methods of image difference, maximum value synthesis and two-dimensional maximum between-class variance. Secondly, the decision tree is constructed by the classification and regression tree (CART) algorithm then automatically extracts the sandy land information of 2014. Finally, spatial overlay analyze the results of change detection with sandy land information, and realize the pattern of automatic recognition on desertification information. The research shows that the accuracy of change detection-CART decision tree pattern is 89.43% to 93%, and the confidence interval is between 85.90% and 98% at 95% confidence level, clearly that the reliability of its accuracy is relatively high. This pattern not only can make full use of the abundant remote sensing information but also can exclude the interference of redundant information. Obviously, the change detection-CART decision tree pattern is one of the effective methods to identify the desertification information, and it will have important application value to the desertification control project.

Key words: desertification; classification and regression tree; regression tree; change detection; automatic recognition