

# 基于形态学属性剖面 and 单类随机森林分类的道路路域新增建筑物提取方法

史忠奎<sup>1</sup> 李培军<sup>1,†</sup> 罗伦<sup>2</sup> 阳柯<sup>2</sup>

1. 北京大学地球与空间科学学院遥感与地理信息系统研究所, 北京 100871; 2. 中国交通通信信息中心, 北京 100011;

† 通信作者, E-mail: pjli@pku.edu.cn

**摘要** 提出一种利用形态学属性剖面 and 单类随机森林分类的道路路域新增建筑物提取方法。用该方法计算路域范围内两时相高分辨率遥感影像的形态学属性剖面, 将得到的形态学属性剖面与光谱特征叠加, 采用改进的单类随机森林分类方法直接提取新增建筑物。以北京市稻香湖地区两时相高分辨率影像作为实验数据, 对比分析该方法与经典两时相直接分类及分类后比较方法的新增建筑物提取精度。结果表明, 综合利用形态学属性剖面 and 光谱特征提取得到的新增建筑物提取精度比仅使用光谱特征的提取精度显著提高, 其中 Kappa 系数提高 15.11%。此外, 该方法提取结果的 Kappa 系数比两时相直接分类方法提高 1.78%, 比分类后比较方法提高 25.15%, 验证了所提方法的有效性。所采用的单类随机森林方法能够有效地处理高维数据, 并可以度量不同特征对分类结果的重要性。

**关键词** 高分辨率遥感影像; 道路路域; 建筑物变化检测; 形态学属性剖面; 单类随机森林  
中图分类号 P237

## A Method for Extraction of Newly-Built Buildings in Road Region Using Morphological Attribute Profiles and One-Class Random Forest

SHI Zhongkui<sup>1</sup>, LI Peijun<sup>1,†</sup>, LUO Lun<sup>2</sup>, YANG Ke<sup>2</sup>

1. Institute of Remote Sensing and Geographical Information System, School of Earth and Space Sciences, Peking University, Beijing 100871; 2. China Transport Telecommunications & Information Center, Beijing 100011;

† Corresponding author, E-mail: pjli@pku.edu.cn

**Abstract** The authors present a method for extraction of newly-built buildings in road-region using morphological attribute profiles and one-class random forest. The morphological attribute profiles are first obtained from bitemporal high-resolution remote sensing images. The morphological attribute profiles obtained and spectral features are then combined to extract newly-built buildings along road-regions using an improved one-class random forest. Bitemporal images of the Daoxiang Lake area in Beijing are used as experimental data to validate the proposed method, by quantitatively comparing with two conventional change detection methods, i.e., direct bitemporal classification and post-classification comparison methods based on support vector machine. The experimental results show that the accuracy of newly-built building extraction from the proposed method (i.e. using combined spectral features and attribute profiles) is significantly higher than that using only the spectral features, with an increase of 15.11% in Kappa. In addition, the Kappa of the proposed method is 1.78% and 25.15% higher than that of the direct bitemporal classification and that of the post-classification comparison. Therefore, the experimental results validate the effectiveness of the proposed method. Advantages of the one-class random forest include capabilities to effectively deal with high-dimensional data and measure the importance of different features used in one-class classification.

**Key words** high-resolution remote sensing image; road-region; building change detection; morphological attribute profiles; one-class random forest

道路路域指道路在建设和运营过程中所改变和影响的地表自然带状区域。这种带状区域既包括道路本身,也包括道路建设和运营期间与道路相互影响、相互作用的自然生态系统相关区域<sup>[1]</sup>。及时准确地掌握道路路域新增建筑物情况,尤其是影响道路交通环境和交通安全的违章建设情况,对于道路维护、道路状况监测以及道路交通安全等具有重要意义。传统的路域新增建筑物监测主要通过人力巡查等方式,耗时长,准确率低,难以进行大范围的动态监测。随着遥感技术的不断发展,高空间分辨率遥感影像逐步成为获取建筑物变化信息的重要数据源。如何从多时相高分辨率遥感影像中高效、准确地提取道路路域建筑物的变化信息,成为交通遥感领域的重要问题之一<sup>[2-4]</sup>。

国内外学者在利用高分辨率遥感影像进行建筑物变化检测方面开展了大量工作,已取得丰富的成果<sup>[2,5-12]</sup>,但专门针对路域新增建筑物提取方法的研究不多。现有研究多采用基于影像光谱特征的变化检测方法,但由于路域新增建筑物与其周围地物(如道路、裸露土壤等)有显著的光谱相似性,仅依靠光谱信息难以有效地区分。由于高空间分辨率影像包含丰富的空间信息(如结构和形状),因此利用光谱和空间信息进行路域新增建筑物的提取,可获得更高的精度。近年来,作为一种表达空间信息的有效方法,形态学属性剖面(morphological attribute profiles)广泛应用于高分辨率遥感影像信息提取、建筑物识别及建筑物变化检测中<sup>[13-21]</sup>。形态学属性剖面突破了传统方法中以像素或固定邻域为处理单元的局限,将处理单元扩展为具有相似属性的所有像素<sup>[15]</sup>,能够得到高分辨率影像的多种属性信息(面积、转动惯量等)。值得注意的是,在应用形态学属性剖面时,通常需要计算影像的多种属性和多个阈值的特征,得到的属性特征往往维数较高<sup>[16]</sup>。因此,在综合利用形态学属性剖面和光谱信息进行变化提取时,需要首先进行特征降维或采用能够处理高维数据的分析方法。

此外,由于道路路域地物类别的多样性和复杂性,采用传统的多类分类方法对两个时相的影像进行分类以提取变化信息时,不同类别的样本很难获取。如果采用单类分类方法,则只需要选取目标类

样本<sup>[3,22]</sup>。对于路域新增建筑物的提取来说,新增建筑物是唯一的目標类。因此,单类分类方法提供了一种可行的提取方法。在现有的单类分类方法(如单类支持向量机<sup>[23]</sup>、支持向量数据描述<sup>[24]</sup>等)中,基于随机森林的单类分类方法<sup>[25]</sup>可以有效地处理高维数据,已应用于手写体识别、医学影像分类等问题中<sup>[25-27]</sup>,取得较高的分类精度。单类随机森林方法通过模拟生成非目标类(负类)样本,将单类分类问题转变为一般的二类分类问题。但是,现有的单类随机森林方法模拟生成的非目标类样本还难以准确地描述负类的真实概率分布情况。

因此,本文提出一种新的路域新增建筑物提取方法,综合利用形态学属性剖面和光谱特征,采用改进的单类随机森林分类方法来提取新增建筑物。

## 1 方法

首先配准研究区域两时相高分辨率影像,并利用现有道路数据得到道路路域范围对影像进行掩膜,得到道路路域范围内的影像,即道路路域以外的影像不参与后续处理。计算路域范围内两时相影像多种属性的属性剖面,将得到的属性特征与光谱特征叠加起来,利用改进的单类随机森林进行分类,提取路域新增的建筑物。此外,计算并分析不同特征对分类结果的重要性。方法流程如图 1 所示。

### 1.1 形态学属性剖面

形态学属性剖面是基于属性滤波完成的。属性滤波是根据某一准则对图像进行的形态学变换,通过评价图像中连通区域的属性,并按一定准则与其对应的属性阈值进行比较,完成对图像的处理。具体地,属性滤波将灰度图像各连通区域的给定属性  $A$  与给定的属性阈值  $\lambda$  进行比较,对于灰度图像中某一连通区域  $C_i$ ,如果其属性满足设定的条件(例如  $A(C_i) > \lambda$ ),则该连通区域保持不变;否则被赋值为邻域相近的灰度值,从而将  $C_i$  合并到周边子块。如果某一连通区域被合并到更低的灰度值区域,则这一操作称为属性薄运算(attribute thinning operation),反之称为属性厚运算(attribute thickening operation)<sup>[14,16,28]</sup>。对于给定的阈值序列  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ ,某一灰度图像  $f$  的属性剖面可以定义为一系列属性厚运算和属性薄运算的组合<sup>[14,28]</sup>。

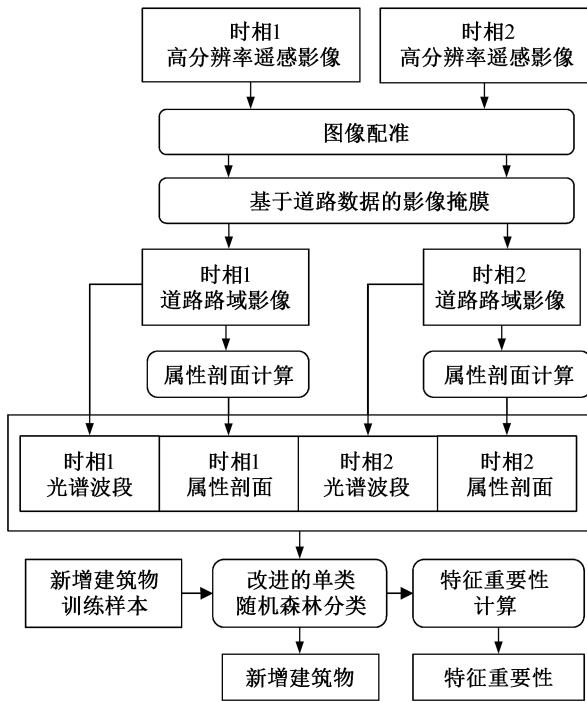


图 1 本文提出的新增建筑物提取方法流程

Fig. 1 Flow chart of the proposed method

$$AP(f) := \{\phi_n(f), \dots, \phi_1(f), f, \gamma_1(f), \dots, \gamma_n(f)\}, \quad (1)$$

其中  $\phi_i$  和  $\gamma_i$  分别表示属性厚运算和属性薄运算。

通过式(1)可以计算得到灰度图像的一种属性剖面。为获得地物的不同属性特征,可以通过不同的属性滤波计算得到多种属性的属性剖面<sup>[14-16,28]</sup>。目前在属性滤波中常用的属性有连通区域面积  $a$ 、连通区域外接矩形对角线长度  $d$ 、连通区域像素灰度值的标准差  $s$  和连通区域的转动惯量  $i$ <sup>[16]</sup>。其中,连通区域的面积  $a$  和外接矩形对角线长度  $d$  衡量连通区域尺寸和形状;灰度值标准差  $s$  描述连通区域内像素灰度值的同质性;转动惯量  $i$  也称为胡氏一阶不变矩<sup>[29]</sup>,描述连通区域的非紧致性,较紧凑的区域转动惯量较低,细长的区域则有着较高的转动惯量<sup>[14,30]</sup>。由式(1)可知,每种属性剖面包含多个阈值的属性滤波结果,因此,计算多种属性的属性剖面时,得到的形态学属性剖面往往特征维数较高。

## 1.2 改进的单类随机森林分类方法

单类随机森林法(one-class random forests)是一种基于随机森林<sup>[31-32]</sup>的单类分类方法<sup>[25-27]</sup>,其单类分类策略是利用目标类(正类)样本的概率分布信息

模拟生成非目标类(负类)样本,进而将单类分类问题转变为二类分类问题<sup>[25]</sup>。

单类随机森林法假设负类概率分布与正类概率分布呈互补关系,利用正类样本的概率分布信息计算得到负类的概率分布函数,并根据这一负类概率分布函数模拟生成人工负类样本<sup>[25]</sup>。在文献[25]中,单类随机森林法生成人工负类样本的策略是,在正类样本分布稀疏的特征空间生成较多的负类样本,在正类样本分布密集的特征空间生成较少的负类样本。文献[25]采用标准化直方图描述正类的概率分布,但这样难以准确地描述正类的高维概率分布,因而由此估计得到的负类概率分布函数也不够准确。依照不准确的负类概率分布函数,可能产生大量带有错误信息的负类样本。

本文提出一种改进的单类随机森林分类方法,所使用的负类样本不是模拟生成,而是在整个图像随机选取的未标记(即没有类别标签)样本中筛选得到。在构建分类器的过程中,除正类样本外,未标记样本同样可以提供有用的信息<sup>[33-34]</sup>。在图像中随机选取的未标记样本包含正类样本和负类样本,如果能够从中筛选出负类样本,就能利用筛选得到的负类样本与已有的正类样本训练二类分类器。本文首先将未标记样本全部视为负类样本,与已有的正类样本一起训练随机森林分类模型;然后使用该随机森林模型对未标记样本进行分类,仅将负类后验概率达到一定阈值(例如 0.8)的未标记样本筛选为负类样本;再将这些筛选得到的负类样本和已知的正类样本一同再次训练新的随机森林模型,对全图做最终分类。之所以在第一步中可以将未标记样本全部视为负类样本,是因为随机森林方法对含有部分错误类别标签的训练样本(在这里即含有部分正类样本的未标记样本)有着较强的鲁棒性<sup>[35]</sup>,少量的错误标记样本对分类结果影响有限(例如文献[35]实验结果表明,在二类分类中,错误标记样本比例达到 25%时,整体精度仅下降 2.8%)。

本文提出的改进的单类随机森林分类方法主要包含 3 个步骤: 1) 从全图随机选取未标记样本,将其全部作为负类样本,并与已知的正类样本一同训练随机森林; 2) 根据随机森林分类得到的样本后验概率,以一定阈值筛选样本,将较高概率为负类的样本筛选为新的负类样本; 3) 将筛选得到的负类样本与已知的正类样本一起再次训练随机森林,得到

最终分类器。

与随机森林分类方法<sup>[31-32]</sup>类似,改进的单类随机森林分类方法同样具有分析特征重要性(feature importance)的能力。在构建分类器时,未被选中的样本组成袋外(OOB)数据集,通过袋外数据集可以统计某一特征对应的平均置换精度差异,从而得到每个特征的重要性指标<sup>[29-30]</sup>。本文计算两个时相不同形态学属性剖面 and 光谱特征的特征重要性,分析比较形态学属性特征与光谱特征在新增建筑物提取中的特征重要性比重。此外,本文还将对比分析不同时相特征的重要性。

### 1.3 精度评价

实验结果的精度评价包含专题精度和几何误差两个方面。专题精度包括总体精度(OA)、Kappa 系数、生产者精度(PA)和用户精度(UA) 4 个指标。考虑到由高空间分辨率影像得到的提取结果往往具有丰富的空间细节,仅使用专题精度难以充分衡量结果的几何特性,因此本文使用几何误差指标进一步评价提取结果的几何精度。本文使用形状指数误差和边界误差两个指标<sup>[36]</sup>。形状指数误差定义为两个对应地物的形状指数差的绝对值,用以衡量提取结果与参考专题图中对应地物的形状差异,两者形状越相似,形状指数误差越小。边界误差指提取结果与参考专题图中对应地物边界重合的误差,两者边界越相似,边界误差越小。

为了验证形态学属性剖面在新增建筑物提取中的作用,分别使用光谱特征、综合形态学属性剖面 and 光谱特征提取新增建筑物,并做对比分析。共进行 5 次交叉验证,取 5 次结果精度平均值作为最终精度值。此外,为验证本文所提方法的有效性,还采用经典的两时相直接分类和分类后比较的变化检测方法提取新增建筑物(这两种方法中均采用支持向量机方法进行分类),将本文方法与这两种方法得到的提取结果进行对比分析。

## 2 实验结果

### 2.1 研究区及实验数据

研究区位于北京市海淀区稻香湖地区。该地区近几年发展迅速,道路路域存在大量新增建筑物。地物类型主要包括建筑物、道路、水泥地面、裸土、植被及水体。本文选取 2012 年 9 月 15 日获取的 GeoEye-1 影像和 2014 年 10 月 18 日获取的 WorldView-2 影像作为实验数据。GeoEye-1 是 2008

年 9 月 6 日发射的高分辨率商业卫星,提供 4 个 1.65 m 分辨率的多光谱波段和 1 个 0.41 m 全色波段影像。WorldView-2 发射于 2009 年 10 月 8 日,提供 8 个 2 m 分辨率的多光谱波段和 1 个 0.5 m 分辨率的全色波段影像。本文采用 Gram-Schmidt 变换法融合全色和多光谱图像,并将融合影像采样到 1 m 分辨率。对两时相的融合图像进行配准,配准的均方根误差为 0.464。按相同范围裁剪图像,最终用图大小为 2000×4000 (图 2)。

本文以影像范围内两条主干道(北清路和稻香湖路)的路域区域作为实验区域。目前,关于路域范围的划定没有统一的标准,研究者往往根据研究对象、目的、内容和方法划定不同的路域范围<sup>[1]</sup>。本文根据道路附近建筑物建设情况,将道路路域设定为从道路外缘起向外 300 m 范围内的区域,如图 2 所示。

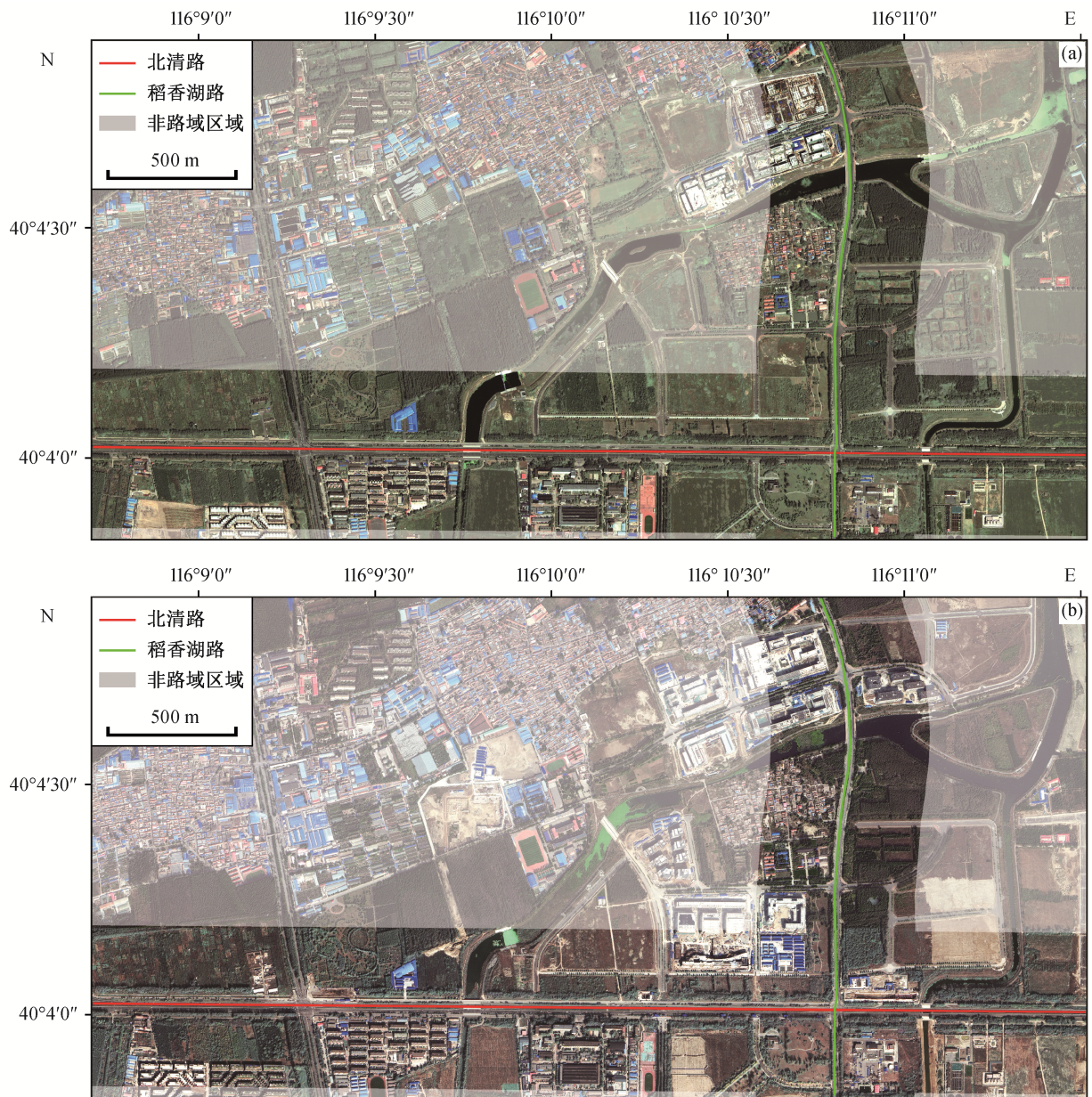
从研究区的影像中可以看出,与时相 1 影像(图 2(a))相比,时相 2 影像(图 2(b))道路两侧新增多处建筑物。新增建筑物大部分空间尺寸较大,具有较亮(水泥材质)或较暗的屋顶(防水材料),也存在少量农村用房(屋顶多为红色砖瓦)及临时用房(蓝色活动板房),空间尺寸较小。本文以具有较大空间尺寸的新增建筑物作为提取目标,通过目视解译得到参考提取结果(图 3),并从中获取训练样本和检验样本。

### 2.2 实验参数设定

本文选取两个时相的全色影像进行形态学属性剖面的计算,其中全色影像同样降采样到 1 m 的分辨率。形态学属性剖面包含前面所述的面积  $a$ 、对角线长度  $d$ 、灰度值标准差  $s$  和转动惯量  $i$  这 4 种属性特征。通过目视方法估计道路路域新增建筑物的属性信息,并为每种属性设定 10 个属性阈值。面积属性的属性范围为[250, 2500],步长设定为 250;对角线长度属性范围为[10, 100],步长为 10;灰度值标准差属性范围为[5, 50],步长为 5;转动惯量属性范围为[0.20, 0.65],步长为 0.05。具体的属性阈值设置如下。

$$\begin{aligned}\lambda_a &= [250, 500, 750, 1000, 1250, 1500, 1750, 2000, \\ &\quad 2250, 2500]; \\ \lambda_d &= [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]; \\ \lambda_s &= [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]; \\ \lambda_i &= [0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, \\ &\quad 0.60, 0.65]。 \end{aligned}$$





(a) GeoEye-1 影像, 2012 年 9 月 15 日获取; (b) WorldView-2 影像, 2014 年 10 月 18 日获取

图 2 北京市稻香湖地区两时相真彩色影像

Fig. 2 True-color image pairs over Daoxiang Lake area, Beijing

按上述属性阈值, 计算得到形态学属性剖面, 并与两时相光谱特征叠加起来, 采用改进的单类随机森林分类方法提取新增建筑物。参照文献[25]的方法, 对改进的单类随机森林进行参数设置, 其中树的个数为 200, 子特征个数为总特征个数的算数平方根加 1。每次分类随机选取 2000 个新增建筑物的样本(约占新增建筑物总数的 5%)作为单类分类的训练样本(正类样本), 参考提取结果(图 3)路域范围

内其余部分作为测试数据。为了尽可能多地筛选出有代表性的负类样本, 实验中随机选取的未标记样本个数不宜过小。通过多次实验尝试, 将未标记样本个数设置为正类样本个数的 30 倍, 即 60000 个。

对比实验采用支持向量机分类方法, 利用形态学属性剖面 and 光谱特征, 采用两时相直接分类和分类后比较两种变化检测方法提取路域新增建筑物。支持向量机分类中采用径向基核函数,  $\gamma$  值设置为

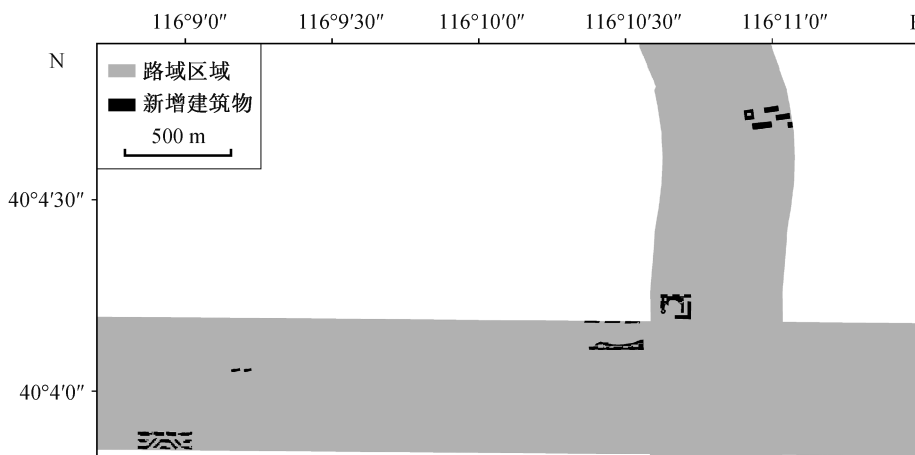


图 3 道路路域范围内新增建筑物参考图

Fig. 3 Reference image of newly-built buildings in road region

特征数目的倒数。其中,两时相直接分类所用目标类训练样本与本文方法所用训练样本相同,非目标类样本则由参考提取结果(图 3)中非新增建筑物区域随机抽取的 60000 个样本点组成。分类后比较的变化检测方法是在两时相影像上分别选取 2000 个建筑物样本和 60000 个非建筑物样本作为训练样本,训练支持向量机模型,分类得到两个时相路域范围内的建筑物,进而比较两时相建筑物结果,得到新增建筑物。

### 2.3 实验结果分析

表 1 显示采用改进的单类随机森林分类,基于光谱特征及综合形态学属性剖面 and 光谱特征的新增建筑物提取结果的精度情况。可以看出,仅使用光谱特征提取的新增建筑物精度较低,Kappa 系数仅为 52.82%,虽然生产者精度高达 73.32%,但是用户精度仅为 42.05%,显著高估了新增建筑物类别。当加入形态学属性剖面(即本文方法)后,提取精度大幅度提高,Kappa 系数提高到 67.93%,生产者精度提升为 81.87%,用户精度也提升为 58.64%。在几何误差方面,综合形态学属性剖面 and 光谱特征的提取结果形状指数误差为 21.77%,边界误差为 22.82%,

优于仅基于光谱的新增建筑物提取结果。综合来看,综合形态学属性剖面 and 光谱特征的提取方法得到的新增建筑物结果具有更高的专题精度和几何精度。

图 4 是采用改进的单类随机森林分类方法提取的新增建筑物结果。可以看出,仅用光谱特征得到的结果(图 4(a))将大量与新增建筑物具有相似光谱特征的地物识别为新增建筑物。例如,在图 4(a)中虚线框位置,大量水泥地面被错误地识别为新增建筑物,显著高估了新增建筑物。综合形态学属性剖面 and 光谱特征的提取结果(图 4(b))中,新增建筑物被准确地提取,建筑物的尺寸和形状信息被准确地保留,大部分与新增建筑物具有相似光谱特征,但具有明显几何属性差异的地物则被正确地识别为非目标类,不过,也存在少量水泥地面被误分为新增建筑物的情况(例如图 4(b)中虚线框位置)。

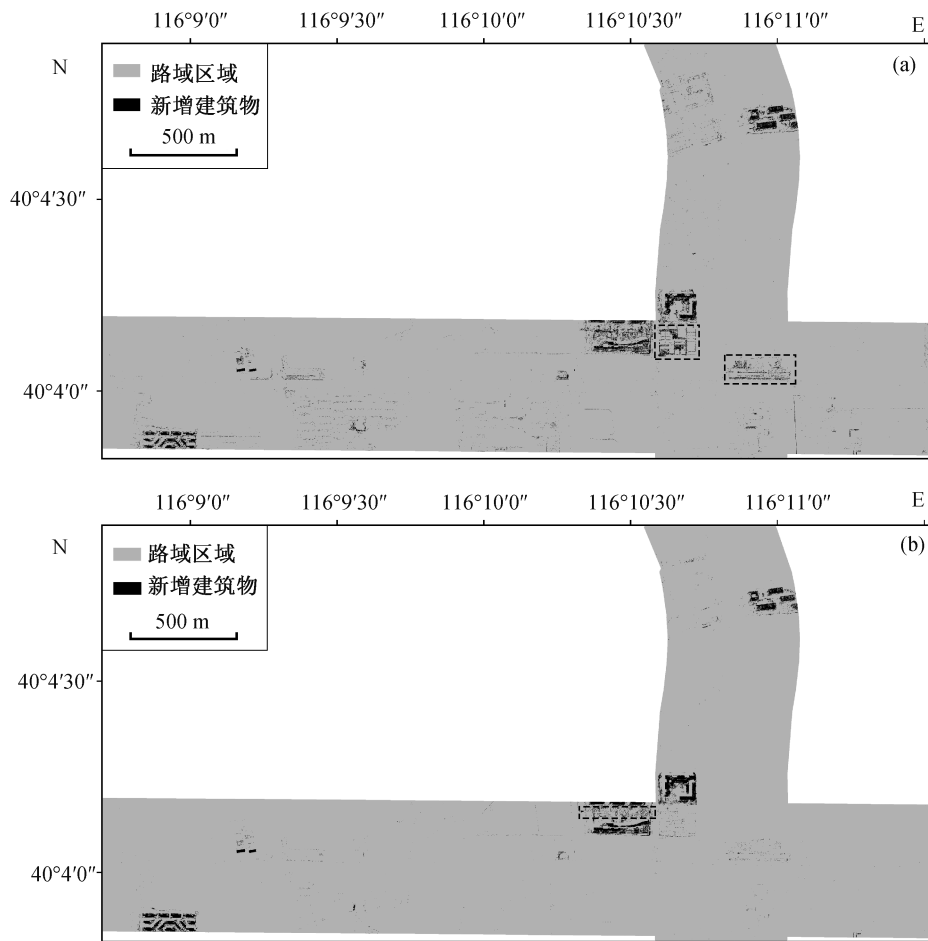
两时相直接分类和分类后比较的变化检测方法新增建筑物提取精度如表 2 所示。可以看出,两时相直接分类得到的新增建筑物提取结果有着较高的专题精度,Kappa 系数为 66.15%,生产者精度和用户精度均达到 60% 以上;但在几何精度方面,边界

表 1 基于不同特征的新增建筑物提取结果精度(%)

Table 1 Accuracy assessment of newly-built building extraction results using different features (%)

特征	专题精度				几何精度	
	OA	Kappa	PA	UA	形状指数误差	边界误差
光谱	98.67	52.82	73.32	42.05	22.81	32.90
光谱+形态学属性剖面(本文方法)	99.21	67.93	81.87	58.64	21.77	22.82

说明:精度为 5 次交叉验证所得结果精度的平均值,分类方法为改进的单类随机森林分类法。



(a) 基于光谱特征的提取结果; (b) 综合形态学属性剖面和光谱特征的提取结果。虚线框为水泥地面误识别为新增建筑物区域。分类方法为改进的单类随机森林分类法

图 4 利用不同特征提取的新增建筑物

Fig. 4 Newly-built building extraction results using different features

表 2 不同方法提取的新增建筑物结果精度(%)

Table 2 Accuracy assessment of newly-built building extraction results obtained by different methods (%)

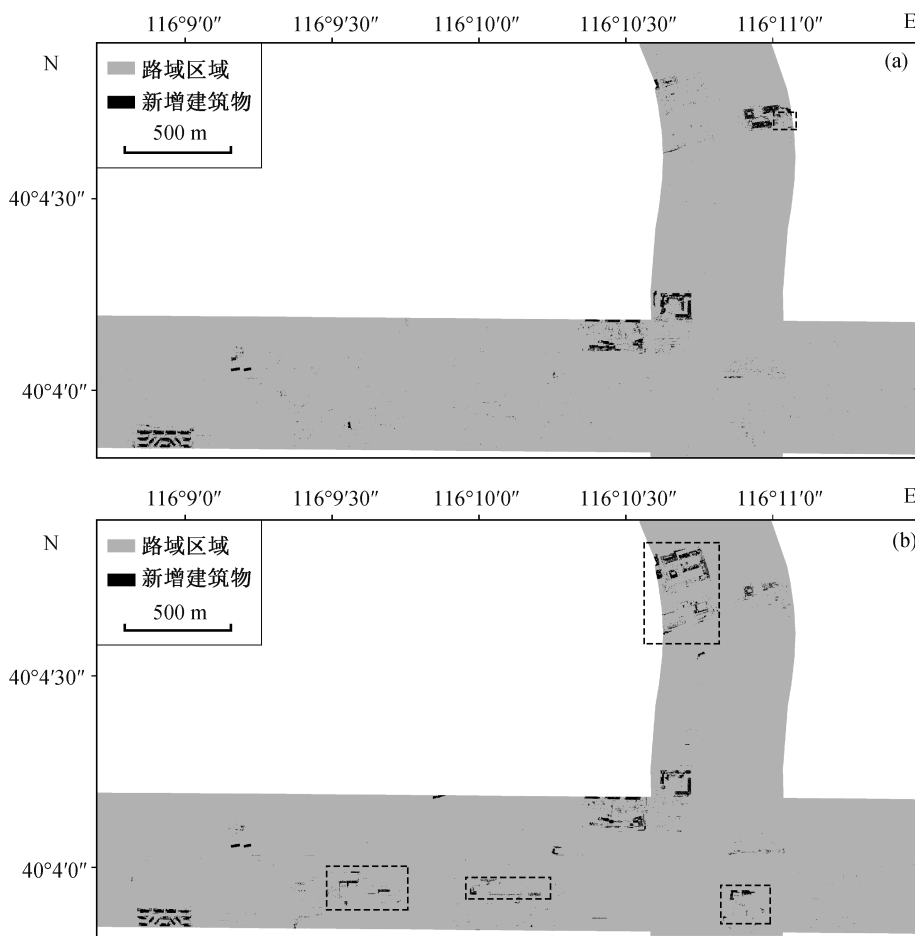
方法	专题精度				几何精度	
	OA	Kappa	PA	UA	形状指数误差	边界误差
两时相直接分类	99.25	66.15	69.30	63.97	19.91	45.65
分类后比较	98.65	42.78	47.36	40.16	15.23	67.26

说明: 所用特征为光谱特征和形态学属性剖面, 分类方法为支持向量机。

误差高达 45.65%。分类后比较的变化检测方法提取得到的新增建筑物结果精度较低, Kappa 系数仅为 42.78%, 生产者精度和用户精度均低于 50%, 边界误差高达 67.26%。将本文方法与这两种方法相比, 本文方法提取结果(表 1)的 Kappa 系数比两时相直接分类方法高 1.78%, 比分类后比较方法高 25.15%; 虽然形状指数误差略高于这两种方法, 但

边界误差比两时相直接分类低 22.83%, 比分类后比较低 44.44%。总体来看, 与传统的两时相直接分类和分类后比较的变化检测方法相比, 本文方法提取结果 Kappa 系数更高, 边界误差更小。

图 5 是两时相直接分类和分类后比较的变化检测方法提取的新增建筑物结果。可以看出, 两时相直接分类较准确地提取了大部分的新增建筑物(图



(a) 两时相直接分类提取结果, 虚线框为新增建筑物漏分区域; (b) 分类后比较提取结果。虚线框为原有建筑物被误识别为新增建筑物区域。采用的特征为光谱和形态学属性剖面, 分类方法为支持向量机

图 5 基于不同方法提取的新增建筑物结果

Fig. 5 Newly-built building extraction results obtained by different methods

5(a)), 但漏分了部分较暗的新增建筑物(图 5(a)虚线框位置)。而分类后比较的变化检测方法虽然较准确地区分了建筑物与非建筑物(图 5(b)), 但将大量原有建筑物混淆为新增建筑物(图 5(b)虚线框位置), 这可能与各时相分类结果的分类误差及两时相建筑物角度差异有关。相比之下, 本文方法可以有效地

区分新增建筑物和原有建筑物(图 4(b))。

表 3 显示不同时相光谱特征和形态学属性剖面的特征重要性。可以看出, 虽然光谱特征对新增建筑物的提取有一定的重要性(15.21%), 但形态学属性剖面的特征重要性更高(84.79%)。此外, 时相 2 所有特征的重要性占总体的 56.43%, 是时相 1 的

表 3 不同时相的光谱特征和形态学属性剖面的特征重要性(%)

Table 3 Importance of spectral features and morphological attribute profiles from images of different dates (%)

时相	光谱	形态学属性剖面				时相汇总
		面积	对角线长度	灰度值标准差	转动惯量	
时相 1	4.14	9.31	9.03	10.18	10.91	43.57
时相 2	11.07	10.14	9.19	10.51	15.52	56.43
特征汇总	15.21			84.79		100.00



1.3 倍,说明时相 2 在新增建筑物的提取中有着更高的特征重要性。

从上述结果可以看出,本文提出的方法比仅用光谱特征的方法和基于支持向量机的方法得到更高的提取精度。一方面,综合利用形态学属性剖面和光谱特征比仅使用光谱特征,得到更高的提取精度,验证了形态学属性剖面在新增建筑物提取中的有效性;另一方面,本文提出的改进的单类随机森林分类方法能够有效地处理高维数据<sup>[19]</sup>,且只利用目标类的训练样本进行训练,得到比支持向量机方法更高的分类精度,同时也简化了样本选择过程。将形态学属性剖面与改进的单类随机森林分类方法有机地结合,是本文方法的一个特点。

需要指出的是,在使用形态学属性剖面时需要设定属性阈值,不同的属性阈值可以产生不同的形态学属性剖面。本文通过目视估计新增建筑物的属性特征设定属性阈值,但由此得到的形态学属性剖面能否有效地表达新增建筑物的属性特征,还需要进一步验证。本文提出的改进的单类随机森林方法能够计算特征重要性,不同属性特征的重要性有可能为属性阈值的选择提供参考,因此,如何利用特征重要性指导属性阈值的设定是未来的研究方向之一。

### 3 结论

本文针对高分辨率遥感影像中路域新增建筑物的特点,将两时相遥感影像光谱特征与形态学属性剖面相结合,采用改进的单类随机森林分类方法,提取路域新增建筑物,并对提取结果进行对比分析。结果表明,与仅用光谱特征的提取结果相比,加入形态学属性剖面可以有效地减少新增建筑物与道路等其他非目标地物的混淆情况,显著改善路域新增建筑物的提取精度,降低提取结果几何误差。分析特征重要性可以发现,在路域新增建筑物提取中,形态学属性剖面对提取结果有着更高的特征重要性。与传统的两时相直接分类和分类后比较变化检测方法相比,本文方法提取结果有着更高的专题精度和更低的几何误差。综上所述,本文提出的方法是一种提取路域新增建筑物的有效方法。进一步的工作是研究更有效的形态学属性剖面的属性阈值设置方法,并进行更多的实验分析,验证本文方法的有效性。

### 参考文献

- [1] 王强. 基于 RS 和 GIS 的山区高速公路生态环境变化检测: 以綦万高速为例[D]. 重庆: 重庆交通大学, 2015
- [2] Huang Xin, Zhang Liangpei, Zhu Tingting. Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2014, 7(1): 105–115
- [3] Li Peijun, Xu Haiqing, Guo Jiancong. Urban building damage detection from very high resolution imagery using OCSVM and spatial features. *International Journal of Remote Sensing*, 2010, 31(13): 3393–3409
- [4] Falco N, Mura M D, Bovolo F, et al. Change detection in VHR images based on morphological attribute profiles. *IEEE Geoscience & Remote Sensing Letters*, 2013, 10(3): 636–640
- [5] 季顺平, 袁修孝. 一种基于阴影检测的建筑物变化检测方法. *遥感学报*, 2007, 11(3): 323–329
- [6] Huertas A, Nevatia R. Detecting changes in aerial views of man-made structures. *Image & Vision Computing*, 2000, 18(8): 583–596
- [7] Stuckens J, Coppin P R, Bauer M E. Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment*, 2000, 71(3): 282–296
- [8] Beumier C, Idrissa M. Building change detection by histogram classification // 17th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS). Dijon: IEEE, 2011: 409–415
- [9] Dong Laigen, Shan Jie. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 84: 85–99
- [10] Tian Jiaojiao, Cui Shiyong, Reinartz P. Building change detection based on satellite stereo imagery and digital surface models. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(1): 406–417
- [11] Vakalopoulou M, Karantzalos K, Komodakis N, et al. Building detection in very high resolution multispectral data with deep learning features // *Geoscience and Remote Sensing Symposium (IGARSS)*. Milan: IEEE, 2015: 1873–1876
- [12] Marin C, Bovolo F, Bruzzone L. Building change detection in multitemporal very high resolution SAR images. *IEEE Transactions on Geoscience and Remote*

- Sensing, 2015, 53(5): 2664–2682
- [13] Soille P. Morphological image analysis, principles and applications. 2nd ed. Berlin: Springer-Verlag, 2003
- [14] Mura M D, Benediktsson J A, Waske B, et al. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience & Remote Sensing*, 2010, 48(10): 3747–3762
- [15] 蒋李兵, 王壮, 郁文贤. 基于属性滤波和上下文分析的高分辨遥感图像建筑物提取方法. *电子与信息学报*, 2012, 34(12): 2985–2991
- [16] Dalla Mura M, Benediktsson J A, Waske B, et al. Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *International Journal of Remote Sensing*, 2010, 31(22): 5975–5991
- [17] Benediktsson J A, Pesaresi M, Amason K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(9): 1940–1949
- [18] Bellens R, Gautama S, Martinez-Fonte L, et al. Improved classification of VHR images of urban areas using directional morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(10): 2803–2813
- [19] Benediktsson J A, Palmason J A, Sveinsson J R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(3): 480–491
- [20] Palmason J, Benediktsson J A, Sveinsson J R, et al. Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis // *IEEE International Geoscience and Remote Sensing Symposium 2005*. Seoul, 2005: 176–179
- [21] Dalla Mura M, Villa A, Benediktsson J A, et al. Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(3): 542–546
- [22] Tax D M J. One-class classification [D]. Delft: Delft University of Technology, 2001
- [23] Manevitz L M, Yousef M, Cristianini N. One-class SVMs for document classification. *Journal of Machine Learning Research*, 2001, 2(1): 139–154
- [24] Srivastava C. Support vector data description. *Machine Learning*, 2004, 54(1): 45–66
- [25] Désir C, Bernard S, Petitjean C, et al. One class random forests. *Pattern Recognition*, 2013, 46(12): 3490–3506
- [26] Désir C, Bernard S, Petitjean C, et al. A random forest based approach for one class classification in medical imaging // *3rd MICCAI International Workshop on Machine Learning in Medical Imaging (MLMI)*. Nice, 2012: 250–257
- [27] Désir C, Bernard S, Petitjean C, et al. A new random forest method for one-class classification // *IAPR International Workshop on Statistical Techniques in Pattern Recognition, SPR 2012*. Notes in Computer Science, Vol. 7626. Hiroshima, 2012: 282–290
- [28] Song Benqin, Li Jun, Dalla Mura M, et al. Remotely sensed image classification using sparse representations of morphological attribute profiles. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 52(8): 5122–5136
- [29] Hu M K. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 1962, 8(2): 179–187
- [30] Gonzalez R C, Woods R E. Digital image processing. 3rd ed. Englewood Cliffs: Prentice-Hall, 2008
- [31] Breiman L. Random forests. *Machine Learning*, 2001, 45(1): 5–32
- [32] Gislason P O, Benediktsson J A, Sveinsson J R. Random forests for land cover classification. *Pattern Recognition Letters*, 2006, 27(4): 294–300
- [33] Elkan C, Noto K. Learning classifiers from only positive and unlabeled data // *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Las Vegas, 2008: 213–220
- [34] Li Wenkai, Guo Qinghua, Elkan C. A positive and unlabeled learning algorithm for one-class classification of remote-sensing data. *IEEE Transactions on Geoscience & Remote Sensing*, 2011, 49(2): 717–725
- [35] Mellor A, Boukir S, Haywood A, et al. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry & Remote Sensing*, 2015, 105: 155–168
- [36] Persello C, Bruzzone L. A novel protocol for accuracy assessment in classification of very high resolution images. *IEEE Transactions on Geoscience & Remote Sensing*, 2010, 48(3): 1232–1244