

分类号：TP79

单位代码：10110

学 号：S20110280

中 北 大 学
硕 士 学 位 论 文

基于决策树和混合像元分解的
城市绿地信息的提取

硕士研究生 王飞红
指导教师 杨建生 副教授
学科专业 测试计量技术及仪器

2014 年 5 月 27 日

基于决策树和混合像元分解的城市绿地信息的提取

王飞红

中北大学

图书分类号 TP79

密级 非密

UDC^{注 1} _____

硕 士 学 位 论 文

基于决策树和混合像元分解的城市绿地信息的提取

王飞红

指导教师(姓名、职称) 杨建生 副教授

申请学位级别 工学硕士

专业名称 测试计量技术及仪器

论文提交日期 2014 年 5 月 27 日

论文答辩日期 2014 年 5 月 22 日

学位授予日期 年 月 日

论文评阅人 王健安 赵冬娥

答辩委员会主席 肖连团

2014 年 5 月 27 日

原创性声明

本人郑重声明：所呈交的学位论文，是本人在指导教师的指导下，独立进行研究所取得的成果。除文中已经注明引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的科研成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律责任由本人承担。

论文作者签名：_____ 日期：_____

关于学位论文使用权的说明

本人完全了解中北大学有关保管、使用学位论文的规定，其中包括：

①学校有权保管、并向有关部门送交学位论文的原件与复印件；②学校可以采用影印、缩印或其它复制手段复制并保存学位论文；③学校可允许学位论文被查阅或借阅；④学校可以学术交流为目的，复制赠送和交换学位论文；⑤学校可以公布学位论文的全部或部分内(保密学位论文在解密后遵守此规定)。

签 名：_____ 日期：_____

导师签名：_____ 日期：_____

基于决策树和混合像元分解的城市绿地信息的提取

摘要

绿地作为城市生态平衡的调控者，在城市生态建设中发挥着重要的功能，除了具有美化环境、净化空气、改善气候等多种生态功能外，它还可作为居民提供娱乐、游览场所，满足居民的精神和文化生活需要。由于绿地可以提高城市景观的自然性、推进城市居民与自然的和谐发展，因此，绿地已成为城市现代化和城市文明的重要标志。近年来，由于我国城市化进程加快，用地矛盾突出，致使绿地急剧减少，导致城市出现了多种诸如“热岛效应”、暴雨、沙尘暴等不和谐自然现象的发生。因此，如何保护绿地，并对绿化区域进行合理分布已成为急需解决的问题。

遥感技术对地物信息获取具有宏观性、周期性、经济性、客观性等优势，为绿地信息的准确获取提供了重要的技术手段。本文应用 TM 遥感影像，利用决策树和混合像元分解相结合的方法对太原市绿地信息进行提取。我们结合水体指数、建筑指数与植被指数，并且综合主成份分析和缨帽变换建立决策树，提取出绿地信息（草地、针叶植被和阔叶植被）；然后使用 MNF 变换和 PPI 端元提纯，应用 N 维散点图进行端元确定，通过线性分解变换得到绿地的丰度值，再乘以单个像元面积，得到绿地的面积。

决策树和混合像元结合应用于绿地信息提取的精度为 87%以上，较决策树提取的精度提高 4%以上，优于常规的分类方法。结果表明：决策树与混合像元分解的方法对分辨率遥感数据信息提取，提高了监测精度，为城市规划提供了客观参考。

关键字：遥感，城市绿地，决策树，混合像元分解，TM 影像，太原市

Study on extraction methods of urban green space information based on decision tree and decomposition of mixed pixels

Abstract

As the regulator of the urban ecological balance, Green space plays an important role in construction of ecological city. It can beautify the environment, clean air and improve the climate. In addition to these ecological functions, it also can provide entertainment and sightseeing places for residents, and meet the needs of the people's spiritual and cultural life. Green space can improve the naturalness of city landscape, and promote the harmonious development of the urban residents and nature, therefore, green space has become an important symbol of urban modernization and civilization. In recent years, China's urbanization process is accelerated, and land contradiction is obvious, resulting in a sharp reduction of green space. These lead to the emergence of a variety of the city problem, such as “urban heat island effect”, heavy rain, dust storms and other disharmonious natural phenomenon. Therefore, there is an urgent problem that how to protect and reasonably distribute the green space.

Remote sensing technology provides important technical means for obtaining accurate green space information on account of its merits in its macro in its cyclical, economic, and objectivity aspects. In this paper, TM remote sensing image is applied, and the combination decision tree and mixed pixel decomposition is used to extract urban green space information. We combine water index, building index ,vegetation index, the principal component analysis and tasseled cap transformation to build decision tree and extract the green space information (grassland,coniferous vegetation and broad-leaved vegetation); And then the MNF transform and PPI are used to purify endmember, and N dimensional scatter plot is applied to determine endmember. Finally, abundance value of green space obtained by linear decomposition

transformation multiply by a single pixel area, we get the green area.

The combination of decision tree and decomposition of mixed pixel is used to extract the green information, and accuracy is above 87%. Compared to decision tree, classification accuracy is improved by more than 4%, showing superior to conventional classification. The results show that: the combination of decision tree and decomposition of mixed pixel is used to obtain information for Moderate-Resolution remote sensing data, which improves the monitoring precision and provides objective reference for urban planning.

Key words: Remote sensing, Urban green space, Design tree, Decomposition of mixed pixels, TM image, Taiyuan city

目录

1 绪论	
1.1 研究背景及其意义	1
1.2 国内外研究现状及发展动态	2
1.2.1 城市绿地调查研究	3
1.2.2 绿地提取研究	3
1.3 研究内容及技术路线	7
1.3.1 研究内容	7
1.3.2 技术路线	8
1.4 本文内容及安排	9
2 研究区简介和数据预处理	
2.1 研究区简介	10
2.2 数据源	12
2.2.1 Landsat TM	12
2.2.2 ZY3 数据	14
2.3 遥感影像预处理	15
2.3.1 几何校正	15
2.3.2 图像裁剪	19
2.3.3 反射率反演	20
2.4 本章小结	22
3 分类特征分析与获取	
3.1 特征介绍	23
3.1.1 主成份变换	23
3.1.2 缨帽变换	26
3.1.3 归一化差异水体指数	27

3.1.4 归一化差异植被指数.....	28
3.1.5 归一化差异建筑指数.....	29
3.2 采样点的特征分析.....	30
3.3 本章小结.....	31
4 基于决策树的绿地信息提取	
4.1 决策树理论.....	32
4.2 绿地提取的决策树模型.....	33
4.3 提取结果的精度评价.....	36
4.4 本章小结.....	38
5 混合像元分解	
5.1 混合像元分解原理.....	39
5.1.1 混合像元分解模型.....	39
5.1.2 线性光谱混合模型.....	40
5.2 混合像元分解模型的建立.....	40
5.2.1 端元组分的确定.....	45
5.2.2 线性混合像元分解.....	46
5.3 精度评价.....	47
6 总结与展望	
6.1 论文总结.....	49
6.2 不足与展望.....	49

参考文献

攻读硕士期间发表的论文

致 谢

1 绪论

1.1 研究背景及其意义

自二十世纪以来,随着科学技术与中国城市规划理论的发展,特别是生态理论在城市规划中的应用,使人们对城市绿地的服务功能的了解,从简单的美化、休闲、娱乐功能,逐步发展到对其生态、防护、教育等综合功能的认识。作为城市整体规划的重要组成部分,城市绿地城市建设中扮演着重要的角色,包括生态,经济和社会效益。首先绿地作为城市生态平衡的调控者,在城市建设中有着重要的生态效益,拥有净化空气、维持碳氧均衡、吸收有害气体、调节和改良小气候、美化环境等多种生态功效;其次,城市绿地具有一定的经济效益,中国的园林艺术是一项具有民族特色和突出优势的重要旅游资源,可以为城市创造可观的经济效益;另外城市绿地不仅有着重要的经济与生态作用,也有着重要的社会与心理效用。绿地有益于城市居民身心的健康发展,可以减轻城市紧张生活带来的压力。另外影响着城市的空间结构,体现在阻隔或引导城市发展,塑造城市发展的结构与形态,对城市的影响可以分为城市物理空间的影响和社会文化空间的影响。最后,城市绿地,可以提供一个舒适和美丽的日常休息、旅游、娱乐和体育休闲活动的场所,让人们到绿地来休憩、游玩、健身、锻炼等,从而获得很好的休息、放松与调节,以消释日常工作的紧张,疲惫,恢复体力,振奋精神,提高工作效率,由此充分体现出了城市绿地的社会效益。因此,一个城市绿地的发展和规划在城市发展中发挥着重要的作用,与城市的发展紧密相联,不可分割,城市绿地面积和分布已经被公认为是城市现代化和城市文明的重要标志。

近年来,工业化和城市人口骤增带来了城市的迅速发展,加快了我国城市化进程,致使用地矛盾突出,绿地快速减少,以致出现了多种诸如“热岛效应”、暴雨、沙尘暴等城市问题的出现。在建设和谐城市的总要求下,作为城市建设的重要组成部分,绿地起着不可忽视的作用,其直接影响着城市的空间形态、功能和发展等多个层面。因此,如何保护绿地,并对绿化区域进行合理分布已成为急需解决的问题。

绿地的保护和合理化区域分布的首要前提是对当下的绿地分布情况和面积占有率情况进行准确调查。调查研究绿地的方法大致分为两种,即人工统计调查和遥感监测调查。两者相比较而言,人工统计方法费时费力,不同的工作人员之间存在个体差异,导致调查结果差异较大,准确度不高,尤其是在大范围区域进行调查工作时,其结果误差较大。同时,该方法所消耗的物力和财力也是非常可观的。而遥感方法使用卫星成像的方式对地表进行监测,其监测范围大,周期性强,重复率高,能够科学、准确、及时地获取第一手资料,能够大量地节约时间、人力和财力,产生很高的社会和经济效益,更为重要的是该方法可以大大降低人为因素的干扰,使监测结果更具有客观性。因此,遥感方法被越来越广泛地应用到城市地物信息提取,尤其是绿地信息提取研究,为城市的合理规划提供了众多的科学依据,并取得了很好的应用效果,如北京市的 8301 工程,上海市的第三轮遥感综合调查,又如广州市、天津市、桂林市都应用遥感影像,制作了城市绿地分布、绿地类型等图像后并用于定量研究等。

从 2005 年开始,各级绿化部门依照构建社会主义和谐社会的要求,以邓小平理论、“三个代表”重要思想为引导,全面落实科学发展观,唱响“以人为本,共建绿色家园”的主旋律,着力提高全社会对植绿、爱绿、护绿、兴绿的认识,大力推动全民进行绿化,大力推进生态工程的建设,大力推进城乡绿化一体化的建设,加强绿色通道建设。因此,通过遥感对城市绿地进行研究是刻不容缓的。太原市作为山西省的省会城市,是全省的经济、政治、文化、教育、信息、科技、交通中心,不断加快的城市化进程,多元化的社会需求,快速扩张的城市空间,变化着的土地利用环境,种种这些突出了一个急需解决的问题,如何把握城市绿地系统建设的现状,合理推进太原的城市绿化建设,以实现太原市城乡一体化建设。因此,本文将应用美国陆地资源卫星的 TM 影像和中国资源遥感卫星三号影像,利用遥感的理论研究绿地的提取,获取绿地的分布信息和面积信息,为政府部门的绿地规划提供及时、可靠的决策依据。

1.2 国内外研究现状及发展动态

长期以来,国内外许多学者在城市绿地信息提取进行了不懈的探索和努力,诸多关于绿地信息提取的方法与模型已经提出并得到广泛应用。

1.2.1 城市绿地调查研究

早期的绿地调查采用的是传统的人工调查方法,通过实地测量和估算的方法统计绿地面积,由各基层绿地部门分层实施并上报到绿化管理部门进行汇总,这种方法不仅耗费人力和财力,而且还需要大量的时间,并且人为因素影响严重,结果的可靠性较差。近年来,随着遥感技术的飞速发展,遥感监测调查方法被越来越广泛地应用于绿化部门,通过目视解译和计算机的数字图像处理等手段提取遥感影像中的绿地信息,结合实测数据和历史数据等对城市绿地进行监测。黄慧萍^[1]等(2004)以面向对象和多尺度分割为主要方法,经过样本多边形对象成员函数创建训练样本区,成功得获得大庆市绿地信息,精准获取和快速更新了城市绿地信息的。陈春林^[2]等(2010)利用 TM 遥感影像对哈尔滨的城市绿地进行信息提取,通过比较四种绿地信息提取方法,得出 TM1、3、4 波段合成的方法能够较为理想的获取城市绿地信息。

1.2.2 绿地提取研究

经过遥感影像识别各类地物目标是遥感技术发展的一个关键环节,不论专题制图、专题信息获取、动态变化监测,还是遥感数据库建立等都离不开遥感分类技术。根据其在不同谱段的空间结构特点、光谱特征和其他信息,遥感影像分类将图像中每个像元点或每个区域按照一定的规则分为不同的类别^{[3][4]}。

1) 传统的分类方法

监督分类和非监督分类是遥感图像非常经典的传统分类方法,也是最主要的传统计算机分类方法。

非监督分类,又被称为聚类分析或点群分析,即在多光谱影像中寻找、定义其自然类似的光谱集群的过程。其不需要人工选择训练样本区,只需要事先设定一定的条件,计算机按照某种规则依照光谱或空间等特征形成集群组,而后,研究者将比较每一个集群组和参考数据,将集群组进行类别划分。长期以来已经发展了近百种不同的自然集群算法,比较经典的有 ISODATA 算法、K-均值算法等。

Thonmas^[5]等(1986)提出了一种多时相比值数据非监督分类方法,同时表明该方

法不仅简单易行，并且结果精度较高。

张治英^[6]等（2003）依据 LANDSAT ETM+影像各波段的光谱特征，选取对植被较为敏感的 2、3、4 波段合成图像进行非监督分类，并且将分类成果图与江宁县江滩钉螺孳生地矢量图叠加进行分析，有效地获取了江宁县江滩钉螺孳生地的植被特征。

监督分类，也称训练分类方法，通过确定的样本来识别其他未知的像元的过程。其主要有两个基本环节：1) 选择训练样本和统计信息；2) 选择合适的分类算法。监督分类的精度最主要与训练样本的选择有关，比较常用的监督分类法有最大似然法、最小距离法。

Labovitz M.L^[7]（1986）根据传统的人工从训练样本中提取时效的局限性,提出了训练样本的自动，半自动提取方法，较好的获取地物分类信息。

何国金^[8]等（2003）利用中分辨率 Landsat TM 遥感卫星影像数据，通过野外实测选取若干个地物训练样本区，并在影像上统计训练样本区的信息，对图像进行监督分类，从而获得地物目标信息。

2) 绿地指数提取法

绿色植被的可见光红波段和近红外波段植被光谱中的最典型的波段，植被在可见光红波段具有强烈吸收作用，而在近红外波段具有较高的反射率，因此，这两个波段对同一植被地物具有完全相反的光谱响应，利用这一点可以建立多种组合的植被指数，如比值植被指数（RVI，Ratio Vegetation Index）与归一化植被指数（NDVI, Normalized Difference Vegetation Index)等，被广泛用于绿色植被提取、植被种类分类和其他植被信息获取等多方面^[9,10,11]。

Huete^[12]等（1988 年）为了克服 NDVI 对土壤背景的敏感的不足，创建土壤调整植被指数（SAVI，Soil Adjusted Vegetation Index），后来在此基础上，进一步发展了转换型土壤调节植被指数（TSAVI，Transformed Soil Adjusted Vegetation Index）。

Qi^[13]等（1994 年）在 SAVI 的基础上又提出了改进型 SAVI(MSAVI, Modify Soil Adjusted Vegetation Index)。这些改进的植被指数在减少土壤背景影响方面取得了非常不错的效果^[14,15]。

然而，不同植被之间常存在明显的季节差异，因此，对植被指数序列进行专题信息研究具有重要意义，在利用遥感数据获取植被指数信息时，应该尽可能最大可能的利用地

物之间的物候历差异。

国内外大量研究表明,利用归一化植被指数(NDVI)、差值植被指数(DVI)和比值植被指数(RVI)等各种指数在不同时相的遥感影像变化规律,并结合各种待提取地物的时间序列变化规律进行研究,可以对地物类别信息很好的识别和提取。

REMBOLD^[16]等(2004)利用 TM/ETM+数据以及 NOAA 10 天的 NDVI 时间序列数据,估计每年的作物面积变化。

卫亚星^[17]等(2008)应用密度分割法,对青海的草场进行等级分割。研究表明,这种草地分级方法可以比较精确地反应青海省的土地覆盖类型、每年的地物变化规律以及标定草场冬春场和夏秋场的位置,此外通过 NDVI 值和鲜草产量模型确定产草量和理论载畜量,有效引导畜牧业生产。

陈燕丽^[18]等(2011)应用 AVHRR 和 MODIS 两种 NDVI 时间序列,做出了草原植被的各月和各年之间的变化特征曲线,对草地进行了分析,并对比了两种 NDVI 序列对平均气温、水汽压与降水量三类气候因子的响应差异,为植被监测研究时 NDVI 时间序列的合理选择提供依据。

3)决策树分类方法

决策树分类方法是以某种分割原则为依据,利用树结构把遥感影像数据分割为特征更加均匀的子集,其以分层分类思想为指导原则。决策树分类方法进行影像地物分类分为三大部分:1)判别函数的生成,其需要训练样本;2)建立决策树的分支,根据不同特征确立分支阈值,然后在各个分支子集中重复建立下层结点和分支;3)形成分类决策树。决策树算法可以处理不同空间尺度数据,具有计算效率高、无需统计假设等优点,因此被广泛的应用在遥感影像分类领域中^[19]。

Belward^[20]等(1987)利用决策树分类方法对农作物进行遥感分类,并与最大似然分类方法进行比较,结果表明,虽然两种分类方法的分类结果精度相差不大,但是决策树分类具有计算速度快的优点,对于多维遥感数据的处理是较好的选择。

Schneider^[21]等(2001)利用 MODIS 数据,采用决策树分类方法对城市区域进行分类,取得了较好的分类结果。

Wardlow^[22]等(2008)以美国中央大平原为研究区域,利用多时相的 MODIS NDVI 数据,采用决策树分类方法对农作物进行分类,总体分类精度高于 80%。

何祺胜^[23]等(2006)对渭干河-库车河三角洲绿洲进行研究,建立决策树分类模型,获得当地盐渍地变化信息,获得了较好的提取结果。

孙艳玲^[24]等(2007)利用中分辨率的 MODIS 遥感卫星数据,以山东省作为研究对象,采用决策树分类方法进行土地利用分类试验,获得了较高的分类精度。

4) 存在的问题与发展方向

绿地面积遥感监测,即对遥感影像进行分类提取,在遥感影像上识别绿地的种植面积。非监督分类方法大都应用在未经过野外调查、缺乏先验知识的情况下,需要根据地物统计出来的光谱特征进行分类,其分类结果的精度往往较差。监督分类法需要经过野外采样获取采样点,建立训练样本,然后计算机根据研究者所提供的样本建立判别函数对影像进行分类,这种方法的分类精度取决于训练样本的选取,而且需要每次分类前都要设定训练区,训练样本的质量和数量对绿地最终识别结果产生重要的影响。正是由于研究者的先验知识的加入,监督分类优于非监督分类在许多研究中已经被证实。

植被指数提取法,需要建立植被指数的时间序列,工作量大,且植被指数受地面复杂背景的干扰较大。此外,阈值的准确度将决定分类结果的精度,然而在阈值确定过程中,人为因素影响较大,因此误差较大。

决策树分类方法是通过建立决策树模型进行地物分类,其精度与决策树分类的节点选取、遥感数据选取、判别函数及所建立的决策树模型具有密切的相关关系。相比传统分类方法,具有简单高效的优势。但是面对存在大量的混合像元的中低空间分辨率的影像数据时,分类结果的总体精度相对较低。

以上各种分类方法都是建立在每一个像元只代表一类地表物质的理论基础上,但是在实际中,尤其是在中低分辨率的遥感影像中,很多像元是由两类乃至多类地物构成的,即,混合像元普遍存在。因此,随着对遥感信息分类方法研究的不断深入以及精度需求的不断提高,混合像元问题是各种分类方法所面临的、难以回避的问题。

5) 混合像元分解

由于遥感影像的像元只能具有一个灰度值,而该灰度值是成像时对应面积区域的地物在该波长上的平均反射值,因此,在很多情况下,一个像元的面积并不能准确地划定为某一确定地物的面积。面对这种情况,最为行之有效的办法就是对像元进行解混,即将一个像元内的实际地物种类和面积进行混合像元分解,得到更为准确的某类地物的面

积信息。

在上个世纪七十年代，混合像元的问题已经开始被人们注意到，从九十年代之后，对混合像元的研究逐步增加，即根据不同地物类型在混合像元中的比例或丰度值，然后确定混合像元的地物类型。混合像元分解的方法是基于各个像元在每一个波段的像元值用于估计像元内各种典型地表物质的百分比。混合像元分解模型有多种分解模型，其中有线性分解模型、神经网络分解模型和模糊分解模型等像元分解模型。其中，由于线性分解模型的简单性及其良好的物理意义，它是目前应用最为广泛，也是研究最为深入的模型。

刘永怀^[25]提出了折半像元分解、匹配像元分解、双邻像元分解、相关像元分解四种混合像元分解方法。

Lu 等应用巴西 Rondonia 地区 TM 遥感数据，选取阴影、土壤、绿地三种类型作为端元组分，利用线性混和分解模型对成熟期的连续植被进行了分类并得到面积信息，实验结果验证了该模型的可行性，能获得较高的分类精度；

岳文泽^[27]等通过优化端元选择方法，对上海中心区域 TM 数据混和像元分解，获得较好的精度，结果表明，此优化方法对城市混合像元分解效果较好。

1.3 研究内容及技术路线

1.3.1 研究内容

本课题以山西省太原市为研究区，选用美国陆地资源卫星 Landsat 5 TM 为主要数据源，采用决策树分类方法和混合像元分解相结合的方法对研究区域进行绿地信息提取，进而对绿地分类，获得太原市绿地分布特征和面积信息，主要研究内容包括：

- 1) 研究地物的光谱、指数等特征因子，建立典型地物的波谱曲线，分析各种地物的波谱特征差异，为绿地面积提取提供特征因子；
- 2) 集成多源信息，分析典型地物的特征因子变化曲线，利用直方图确定各种特征的阈值，并运用决策树分类算法，建立决策树分类模型，进而实现研究区域的遥感影像的分类，获取绿地分布范围；
- 3) 研究混合像元分解，包括端元提取算法，确定端元组分，建立混合像元分解模

型，进一步对提取的绿地进行混合像元分解，获取绿地丰度图，从而获得更为精确的绿地面积；

4) 利用实地采集的数据及其它信息，对提取的绿地进行准精度评价，检验研究方法的有效性。

1.3.2 技术路线

为了完成对太原市绿地的提取，本论文主要分为中分辨率遥感影像数据获取、遥感影像数据预处理以及影像信息提取三部分。其中遥感影像信息的提取是关键，直接决定了绿地提取的精度和可靠性。技术路线如下图 1.1 所示：

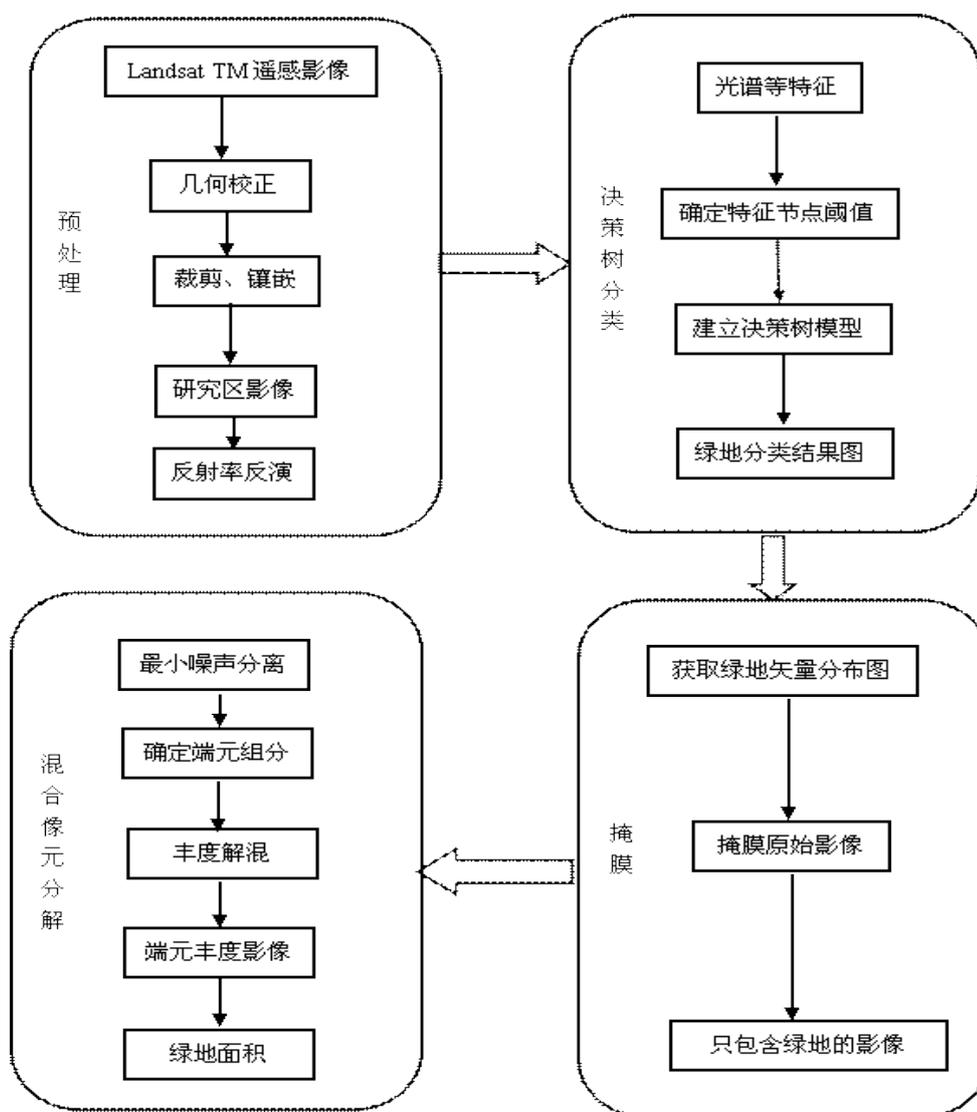


图 1.1 本论文的技术路线图

1.4 本文内容及安排

第一章是绪论，介绍论文的研究背景及意义，概述遥感在绿地信息获取中的国内外研究现状，并简要介绍本文的主要研究内容。

第二章是研究区简介与数据预处理。首先简要概述了研究区域的基本情况及绿地的功能和发展，然后介绍研究中所使用的遥感数据，最后对研究区影像进行了数据的预处理，依次分别为几何校正、图像裁剪和反射率反演。

第三章是采样点的特征分析。介绍了主成份分析、缨帽变换、归一化差异水体指数、归一化差异建筑指数和归一化差异植指数的特征，并对采样点的特征进行分析，作出地物的波谱特征曲线。

第四章是基于决策树的绿地信息提取以及精度评价。介绍了决策树的理论，并构建了绿地提取的决策树模型。

第五章是混合像元分解以及精度评价。介绍混合像元分解的几种模型，阐述了端元提取以及端元确定的方法，建立混合像元分解模型。

第六章是总结和展望。

2 研究区简介和数据预处理

2.1 研究区简介

1) 基本情况

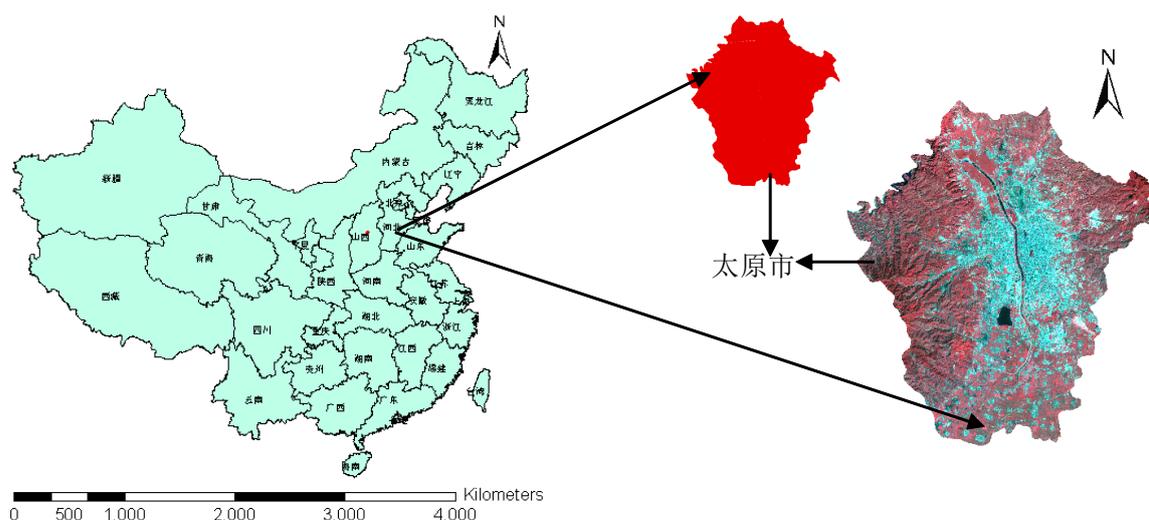


图 2.1 研究区域地理位置

太原市是中国华北地区最重要的城市之一，是山西省的政治、经济、文化中心，濒临汾河，三面环山，也是中国最重要的能源、重工业基地之一。太原市位于山西省的中部，跨越东经 $111^{\circ}30' \sim 113^{\circ}09'$ ，北纬 $37^{\circ}27' \sim 38^{\circ}25'$ ，其东、西、北三面环山，地势南高北低，呈簸箕状。东部山地为太行山余脉，主峰罕山海拔 1591 米，西部山地为吕梁山云中山的延续，主峰北山顶峰 2659 米，为本境最高峰。东、西山之间是太原盆地，海拔高度约 800 米，地势平坦，土地肥饶。其总面积为 1417 平方千米，总人口约 260 万，现辖六区（迎泽区、杏花岭区、万柏林区、尖草坪区、晋源区、小店区）15 个乡镇 48 个街道办事处。居民由汉族、满族、蒙古族、回族、藏族等组成，其中，以汉族为主。汾河贯穿南北，自古以来，有着“控带山河，踞天下之肩背”的气势。

2) 气候情况

太原市位于中国大陆，距离东部海岸线较远，其西北部濒临欧亚大陆腹地，其所处的地理位置和山西高原的地理环境，决定其接收较强的太阳辐射和拥有丰富的光能热

量，在农业上具有较高的光能效率。此外受高太阳辐射和西风环流的影响，导致其气候干燥，白天和晚上的温差较大，降雨量低，年平均降雨量 431.1 毫米。综合以上的地理环境、太阳辐射和大气环流三个方面的综合影响，决定了太原市是典型的大陆性气候，属于温带半干旱季风性气候，夏无酷暑，冬无严冬。春季经常刮风，有较多沙尘；与相同纬度的华北平原地区相比，夏季比较凉爽；秋季天高气爽，晴朗舒服，但较为干燥；冬季较为寒冷。其年平均气温为 10℃左右，存在较长的无霜期，年均约 202 天。太原地区地形复杂，使其形成了差异显著的气候区域，不仅具有明显的垂直变化，并表现出一定程度的水平差异。

3) 绿地功能及发展概况

城市绿地主要是指自然植被和人工植被为主的城市用地，是城市生态建设和发展的主要组成部分，是城市环境质量和居民生活水平的重要标志。绿地具有美化城市景观、改善城市生态环境、塑造城市特色和怡养居民身心的功能，能够为人们创造一个美丽、自然、温馨、和谐的大家园，降低城市生活的心理压力，使得生活在钢筋混凝土高楼林立中的现代城市居民，欣赏到大自然美景，体会到大自然的宜人气息，感受到大自然的巨大力量。

自从改革开放之后，我国的城市绿地发展进程飞速发展，尤其是在这十年中，我国的城区绿地建设高速发展，我国现在城市绿化覆盖率为 28.15%，绿地率达到 23.67%，人均城市绿地达到 6.83 平方米。当然，太原的城市绿化也在高速进行。

2008 年底，太原的人均公共绿地面积 9.6 平方米，与处于同一梯队的石家庄、呼和浩特、昆明、西安等城市相比，在绿地发展方面还有一定差距；2009 年太原市加快绿地建设步伐，跻身“国家园林城市”行列；2013 年一如重视绿地发展，建成区共新增 275 万余平方米的绿地，相当于近 7 个“龙潭公园”（面积约四十万平方米）。在 2013 年太原市园林绿化工程新增的绿地中，213 万平方米城市绿地是园林绿化重点工程增加的，37 万平方米城市绿地是社会绿化增加的，25 万平方米城市绿地是各城区增家的。至此，太原市的绿地覆盖率增加了 0.85%，达到 39.83%。可见，太原市区的绿化近年取得可观的成绩。

2.2 数据源

研究所选用的遥感影像主要为被广泛使用的美国 LANDSAT-TM 遥感卫星影像。辅助数据有：a) 中国 ZY3 高分辨率影像，作为参考影像；b) 1:250000 山西省行政矢量图，用于对研究区影像进行边界剪裁。

研究中使用美国陆地资源卫星 Landsat TM 多光谱遥感影像数据，轨道号 (path/row) 为 125/34，获取时间为 2011 年 4 月 19 日，其投影坐标为 WGS_1984_UTM_Zone_49N，全景影像没有条带和云，成像质量良好。参考数据选择的是 2013 年 4 月 10 日获取的分辨率为 6 米的 ZY3 影像以及 Google Earth 上的高分辨率影像。

2.2.1 Landsat TM

Landsat 是美国 NASA 发射的系列陆地资源卫星，从 1972 年开始，美国发射的第一颗真正的地球观测卫星 Landsat 1，到现在为止，Landsat 计划已经发射了 1 号至 8 号卫星。目前，仍在服务的是 landsat 5、landsat7 和 landsat 8。这里只对 landsat 5 进行详细介绍。

Landsat 5 号卫星于 1984 年 3 月发射，其工作状态稳定良好，地球影像的获得几近实现了连续。其采用高度 705 千米、轨道倾角 98.2 度的太阳同步准回归轨道。每 16 天扫描同一地区，即 16 天覆盖全球一次，地方平均时 9: 39AM 通过赤道上方，扫描宽带为 185 千米。Landsat 5 号卫星上搭载 TM (Thematic Mapper) 专题制图仪。

TM 影像为多光谱遥感影像，共包含有 7 个波段，第 1-5 波段和第 7 波段的空间分辨率为 30 米，它们对地物辐射信息的接收灵敏度相对较高，波段 6 的空间分辨率为 120 米，其为热红外波段。TM 影像有着丰富的光谱信息和较高的更新频率，其运行时间长且数据时间上完整，因此被广泛应用于遥感图像处理。TM 影像的各个波段的具体参数参照表 2.1。本文应用的 TM 遥感影像数据如下图 2.2 所示。

表 2.1 Landsat 5 TM 各波段的特征参数

波段	光谱分辨率/ μm	空间分辨率/m	主要作用
1	0.45~0.52 (蓝)	30	用于判别水深、水中泥沙分布, 分辨土壤植被
2	0.52~0.60 (绿)	30	用于评价绿色植物的生活力, 分辨植被, 反映水下特征
3	0.63~0.69 (红)	30	处于叶绿素强吸收区, 用于观测道路/裸露土壤/植被种类
4	0.76~0.90 (近红外)	30	用于生物量调查/作物长势测定/农作物估产, 辨水体和潮壤
5	1.55~1.75 (中红外)	30	用于分辨道路/裸土/水, 在不同植被之间有良好的对比度, 并有较好的穿透大气、云雾的能力
6	10.4~12.5 (热红外)	120	根据辐射响应的差别, 区分与人类活动有关的热特征
7	2.08~2.35 (中红外)	30	用于分辨岩石/矿物和辨识植被覆盖和湿润土壤

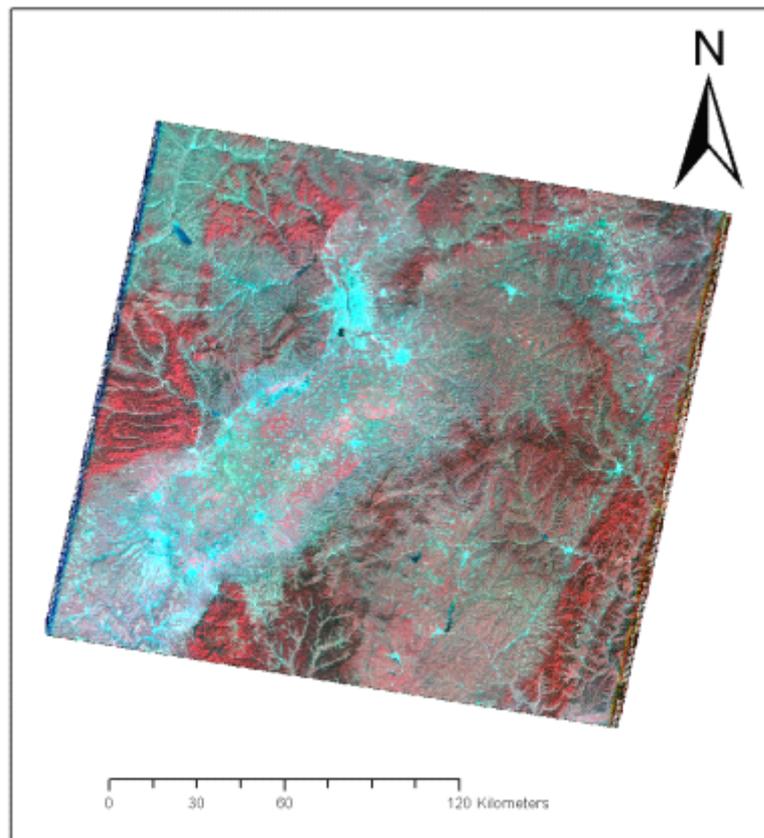


图 2.2 TM 原始影像数据 (RGB=5, 4, 3)

2.2.2 ZY3 数据

资源三号卫星（简称 ZY3）是我国的首颗高分辨率的民用卫星，于 2012 年 1 月成功发射，轨道高度为 505.984 千米，轨道倾角为 97.421°，近地点幅角是 90°，且采用的是与太阳同步的准回归轨道。在赤道正上方时，地方平均时是上午 10 点半。其回归周期为 59 天，侧翼的重访时间是 3 到 5 天左右。ZY3 卫星上提供了覆盖全国的多光谱影像和高分辨率立体影像，服务于国土资源的监测及城市规划等领域。

ZY3 卫星搭载着前视相机、后视相机、正视相机和多光谱相机，其中，正、前、后相机能够获取同一地区的不同观测角度的立体像对，可以提供丰富的三维信息，为我国的立体测图这一领域提供了方便。多光谱相机不仅具有较为丰富的光谱信息，其一共包含四个波段，而且具有较高的空间分辨率，其为 6 米，这些都有利于研究中地表物质的识别。表 2.2 是 ZY3 的主要参数。

研究中使用的 ZY3 遥感影像为 1A 级产品数据，因此首先对其进行正射校正，应用 ENVI 中的 QuickBird RPC 校正模型对影像进行校正，从而实现与 TM 影像的投影一致，获得的 ZY3 影像如图 2.3 所示。

表 2.2 ZY3 的主要参数

有效载荷	波段号	光谱范围(μm)	空间分辨率(m)	幅宽(Km)	重访时间(天)
前视相机	—	0.50~0.80	3.5	52	3~5
后视相机	—	0.50~0.80	3.5	52	3~5
正视相机	—	0.50~0.80	2.1	51	3~5
多光谱相机	1	0.45~0.52	6	51	5
	2	0.52~0.59			
	3	0.63~0.69			
	4	0.77~0.89			

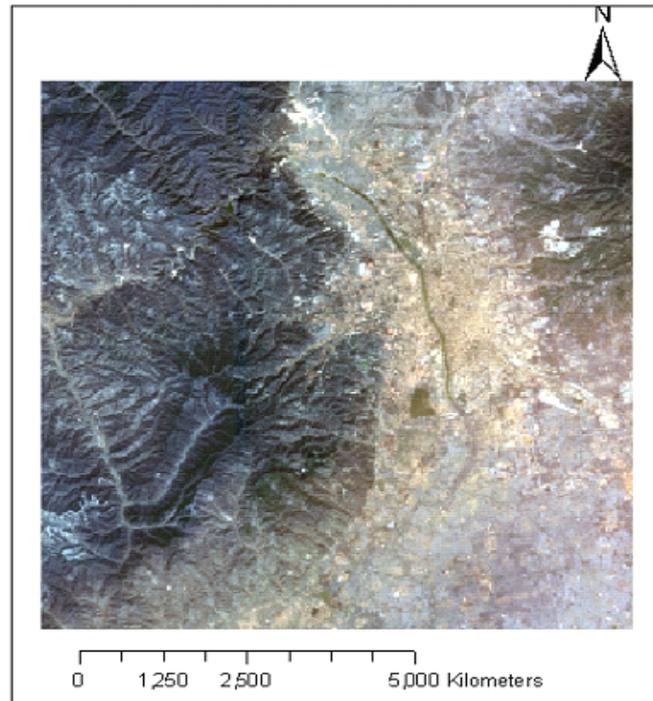


图 2.3 ZY3 卫星影像数据

2.3 遥感影像预处理

2.3.1 几何校正

由于遥感卫星成像时，由于地球自转、遥感平台的位置和运动状态的变化、地球表面曲率、传感器等因素的影响，使得卫星图像与实际分布情况存在差异，导致产生畸变。由于畸变的存在，会使成像影像相对实际地物出现形状的伸展与扭曲，位置的偏移与挤压等现象，这些现象都会对遥感图像的信息提取和分析形成一定的阻碍，所以，在使用遥感数据前，首要的是对数据进行位置校正，即“几何校正”。

几何校正包括几何粗校正和几何精校正两种。前者一般有专门的系统几何校正的软件包，用于校正引起几何畸变的原因，在遥感数据处理中心完成；后者是利用包含精确经纬度信息的地面控制点的卫星影像或地理信息（校正参考标准）进行的几何校正。为了使遥感图像和实际地物拥有一致的空间位置与对应的光谱特征，几何精校正是必需的。精校正多利用数学建模的方式进行，即在遥感图像中根据地理位置信息固定的地物，寻找其在参考标准中的信息，将遥感影像中的经纬度信息与参考标准中的经纬度信息确

定数学关系，进而建立两者之间的数学模型，并利用该模型计算整幅遥感图像各像元的位置并重新进行排布。校正过程中，还需要注意的一个问题是像元的赋值，这是因为像元在几何校正过程中不可避免地要发生位移，导致像元面积的扩充、压缩或消失，在一定程度上改变了原有的信息。因此，几何精校正过程中需要有合适的方法来保证图像的原始信息，这种方法称为像元重采样。在研究中，所使用的影像数据均为二级产品，只经过了系统级的几何粗校正，因此需对影像进行几何精校正。本文在进行几何校正时，采用高分辨率遥感数据作为参考影像，具体操作步骤如下，流程图如 2.4 所示。

1) 选取地面控制点

在待校正图像选取典型地物点，同时相应地在参考图像（已经过校正的图像）上选取准确的地面控制点（GCP），这样就获取了地面控制点对，重复选取足够的 GCP 点对。为了使几何精校正精度更为精确，我们应该在待校正的图像上和已校正图像上选取地理信息固定的点为采样点，包括比较清晰的、典型的地物目标的线状地物的交角或地物的拐角处，如道路交叉处、大型建筑物的顶端和花坛的边缘等；其次，在已校正图像和待校正图像上选取 GCP 时，不要选择随时间而变化的地物点，如林地的边界和河流的拐角；最后，GCP 的分布在图像中是均匀的，且在图像的四角周围都要选取一个地物点，同时影像中山地地形变化比较频繁的地方畸变较为严重，这些区域的 GCP 密度较整幅图像平均 GCP 密度略高。完成 GCP 点对的选取后，验证控制点对的成图精度，去除偏差太大的控制点对，同时更新控制点对，重新选择更多的 GCP 点对，最终确保确保控制点对的数量和质量，保证较高的成图精度。

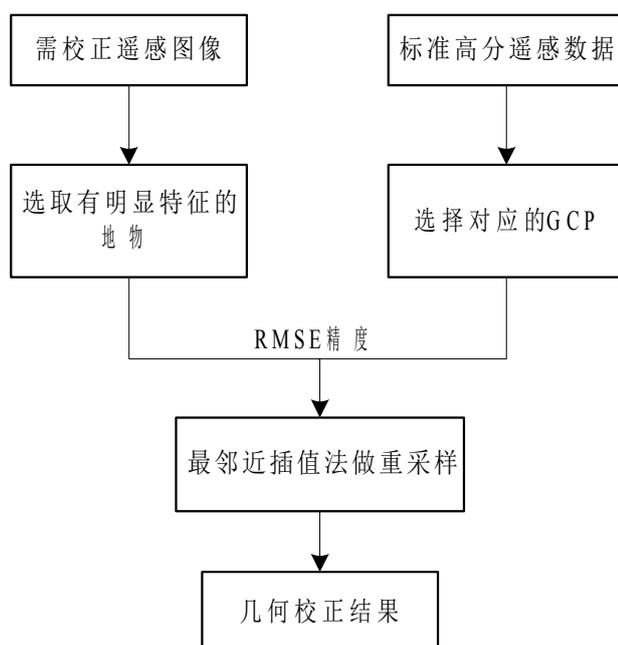


图 2.4 几何精校正流程图

2) 确定几何校正模型

完成对 GCP 点对的选取后，需要确定 GCP 点对空间变换函数，鉴于多项式变换模型具有较高的认可，我们将其确定为几何校正的数学模型，其没有考虑图像失真的原因，根据两幅图像中形成的 GCP 点对建立两者之间的关系^[28]。其表达式为：

$$\begin{aligned}
 x &= a_0 + (a_1X + a_2Y) + (a_3X^2 + a_4XY + a_5Y^2) + (a_6X^3 + a_7X^2Y + a_8XY^2 + a_9Y^3) + \dots \\
 y &= b_0 + (b_1X + b_2Y) + (b_3X^2 + b_4XY + b_5Y^2) + (b_6X^3 + b_7X^2Y + b_8XY^2 + b_9Y^3) + \dots
 \end{aligned}
 \tag{2.1}$$

式中， x, y 为待校正影像的像元坐标； X, Y 为已校正影像的像元坐标。 a_i, b_j ($i, j = 0, 1, 2, \dots, (N-1)$) 为多项式系数，多项式的项数 T 与它的阶数 t 具有以下的关系^[29]：

$$T = \frac{1}{2}(t+1)(t+2)
 \tag{2.2}$$

这里需要注意，在选取 GCP 时，其数目必须大于多项式的系数个数。由于在多项式校正模型的次方数达到 3 以上时，精度计算几乎没有变化，为此，最终选择了三次多项式模型进行几何精校正。根据公式 2.2，控制点应该至少选择 10 个。鉴于研究区处于丘陵地区，为了使校正结果更为精确，最终的 GCP 控制点数量为 30。

3) 图像重采样

重采样是将待校正的图像的像元值进行计算以重新产生一幅校正的图像的处理过程。重采样一般有三种卷积内插法、双线性内插法和最邻近内插法三种方法。三次卷积内插法，其比较复杂，以内插点周围的 16 个像元值进行三次卷积，该方法得到的处理结果具有均衡化和清晰化的效果，然而会改变原有像元的灰度值，此外与后两者比较，其计算量较大^[30]；双线性内插法，根据距离内插点的大小，给予与像元相邻四点的像元不同的权重，从而进行线性内插，该方法得到的处理结果较为平滑，但也会破坏原来的像元灰度值；最邻近内插法的新像元值是其最邻近像元的灰度值，其得到的校正的图像灰度值与原来图像可在最大程度上保持一致。

为了最大程度地保证信息的原始性，本文采用最邻近内插法进行重采样。我们对重采样通过图示(图 2.5)解译其原理，在图中， (x, y) 、 (a, b) 的距离为 1，即 $\Delta x = 1$ ， $\Delta y = 1$ ，通过对比 (x, y) 与邻近四点的距离，显然 (a, b) 点与 (x, y) 点的距离最近，因此，点 (a, b) 的灰度值可以作为 (x, y) 的灰度值。

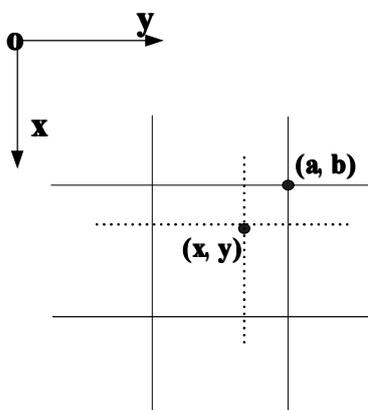


图 2.5 最邻近法图

图 2.6a 是校正标准影像和未校正影像层叠的局部结果，图 2.6b 为校正标准影像和校正后影像层叠的局部结果，可以看出，经几何精校正后，原始图像得到了较好的处理（标识圈内为两个层叠影像中的汾河交汇处），与参考影像的地表物质达到很好的吻合，最明显的就是汾河的交汇处达到一致。

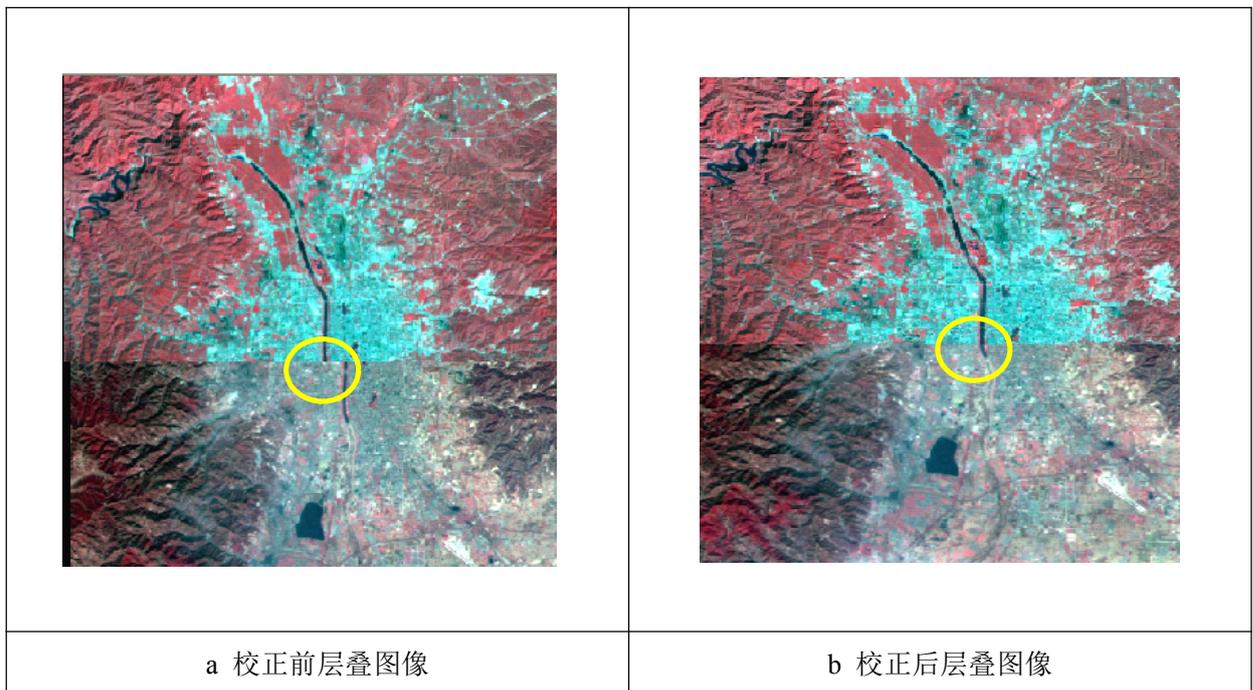


图 2.6 几何校正前后对比图像

2.3.2 图像裁剪

图像裁剪的目的是去除研究之外的区域，获得研究区域^[31]。图像裁剪一般通过两种方法获得感兴趣区（ROI，Region of Interest），一种是手工绘制感兴趣区，另一种是矢量数据生成感兴趣区。

由于太原市市辖区是一景 TM 影像的一部分，为此需要对影像进行剪切获得研究区域。研究中通过矢量数据生成感兴趣区，应用 1: 25 万基础地理信息数据，通过选取和拼接获取太原市行政矢量边界，获得研究区范围的矢量数据，利用其做出 ROI 文件。利用 ENVI 中“subset data via ROI”裁剪图像，获得研究区域，太原市辖区。裁剪得到的研究区域的 TM 遥感影像如图 2.8 所示：

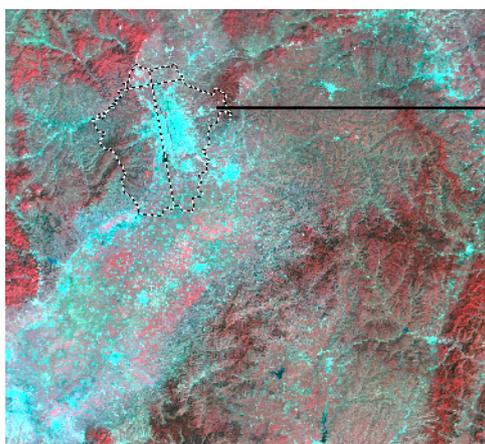


图 2.7 裁剪前图像

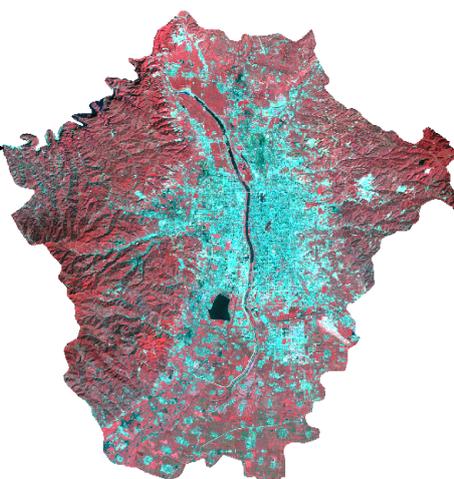


图 2.8 裁剪后图像

图 2.7 为包含研究区域的一景 TM 遥感影像，图中虚线即为研究区域的范围，通过裁剪得到研究区域的 TM 遥感影像如图 2.8 所示。

2.3.3 反射率反演

遥感影像由于受到成像时间、成像角度等因素的影响，存在着“同谱异物”和“同物异谱”的现象。如果以遥感影像中各波段图像的灰度值（DN 值）进行各种运算，势必会导致分类出现错误。对于地物本身而言，由于本身理化特性的不同，反射率在各种波段是不同的，然而反射率基本不会因为光照时间或光照角度而发生变化，因此，遥感影像分类前需要进行反射率反演，在一定程度上可以消除以上两种现象的影响。为了得到更为严密的基础数据，将 TM 光谱值影像转换为反射率影像。其具体工作过程如下：

1) DN 值转换为辐射亮度值

$$R = DN \times Gain + Bias \quad (2.3)$$

式中， R 为像元的辐射亮度值 ($W/(m^2 \cdot sr \cdot \mu m)$)； DN 为原始影像中的像元灰度值； $Gain$ 为对应波段的增益数据 ($W/(m^2 \cdot sr \cdot \mu m)$)； $Bias$ 为对应波段的偏置数据 ($W/(m^2 \cdot sr \cdot \mu m)$)。在 USGS LANDSAT-TM 用户手册中查找 $Gain$ 和 $Bias$ 数据(见表 2.3)。

表 2.3 LANDSAT-TM 的增益数据和偏置数据

<i>Band</i> (波段)	增益数据 $Gain/W / (m^2 \cdot sr \cdot \mu m)$	偏置数据 $Bias/W / (m^2 \cdot sr \cdot \mu m)$
1	0.7268	-1.52
2	1.4425	-2.84
3	1.0399	-1.17
4	0.8726	-1.51
5	0.1199	-0.37
6	0.0551	1.2378
7	0.0653	-0.15

2) 反射率计算

$$\rho_{\lambda} = \frac{\pi \cdot L_{\lambda} \cdot d^2}{E_{SUN_{\lambda}} \cdot \cos\theta_s} \quad (2.4)$$

式 2.4 中， λ 为任一波段的中心波长； d 为日地天文距离（通常取 1）， ρ_{λ} 是第 λ 波段的反射率； $E_{SUN_{\lambda}}$ 是大气顶层太阳辐射平均值，其各个波段的具体参数如表 2.4 所示； L_{λ} 是在第 λ 波段的像元辐射亮度值； θ_s 为太阳天顶角。

计算太阳天顶角的公式为：

$$\theta_s = 90^{\circ} - SUNELEVATION \quad (2.5)$$

式中， $SUNELEVATION$ 是太阳高度角，可以在对应影像头文件中查找。

表 2.4 大气顶层平均太阳辐照度

Band(波段)	1	2	3	4	5	7
ESUN (大气顶层平均 太阳辐照度)/ $W/(m^2 \cdot \mu m)$	1 957	1 829	1 557	1 047	219.3	74.52

根据以上公式，利用 ERDAS IMAGINE 遥感图像处理软件的 MODLE MAKER 建立相应计算模型，最终得到经过反射率反演的影像（图 2.9）。

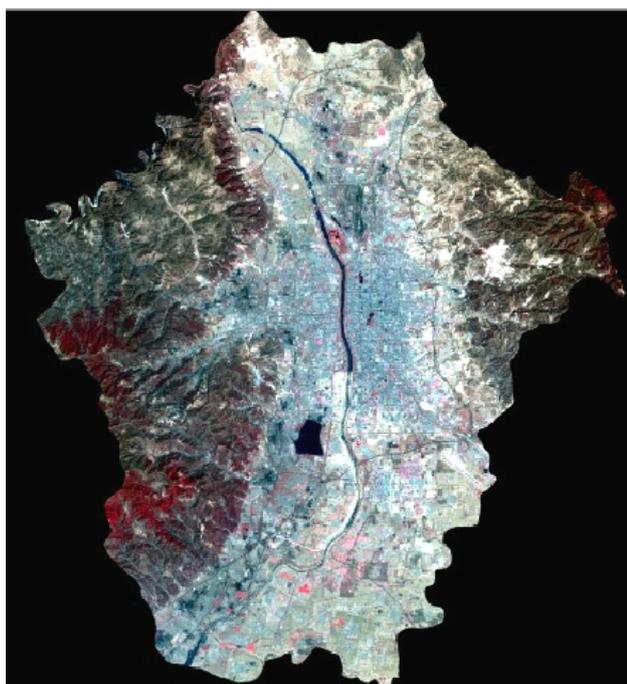


图 2.9 反射率反演影像结果图

2.4 本章小结

本章对研究中所使用到的数据和数据预处理内容进行了介绍。根据研究需要，本文使用了多波段遥感影像数据 TM 作为主数据源，使用高分辨率遥感影像 ZY3、Google 地图、实地采集数据作为参考数据。其次，对研究区数据进行了预处理，首先根据研究区卫星轨道号对遥感影像进行了选取，然后进行几何精校正，最后结合研究区 1:250000 基础地理信息数据完成了剪裁，得到了研究区影像。此外，为了后续指数计算等内容的需要，利用 ERDAS 建模得到了辐射校正结果和反射率反演结果，为后续工作打下基础。

3 分类特征分析与获取

3.1 特征介绍

3.1.1 主成份变换

遥感多光谱影像的波段多，信息量大，在图像处理时，常常耗费大量的时间和占据很大的磁盘空间。实际上，遥感数据各波段之间有一定程度的相关性，存在着数据冗余。主成分分析法（PCA）是一种去除冗余信息，将多波段的图像信息进行压缩，得到比原有波段更有效的少数几个波段的方法，达到保留主要信息、降低数据量和增加类别的可分性的目的^[32]。变换的结果是使得多光谱空间的坐标系统按照某种规律旋转，实现遥感图像线性变换。

主成份变换（PCA）是一种线性变换，是建立在统计特征基础上的变换。主成份变换是利用坐标轴的旋转使得数据的方差达到最大，通过变换矩阵进行线性组合，进而输出互不相关的波段，即重新产生一景具有相同波段的多光谱影像^[33]。其变换的表达式如下：

$$\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} \quad (3.1)$$

式中， \mathbf{X} 是原始影像像的 DN 值， \mathbf{Y} 是新生成的影像 DN 值， \mathbf{A} 为 $n \times n$ 阶的变换矩阵，是 \mathbf{X} 空间协方差矩阵的特征向量矩阵的转置矩阵，它的作用是作为多光谱的像元亮度加权系数，从而实现线性变换。

根据主成份变换的原理可知， \mathbf{A} 是其空间协方差矩阵的转置矩阵， Φ 是其特征向量矩阵，具体表达式如下：

$$\mathbf{A} = \Phi^T = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \Lambda & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \Lambda & \varphi_{2n} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \varphi_{n1} & \varphi_{n2} & \Lambda & \varphi_{n3} \end{bmatrix} \quad (3.2)$$

可以写成

$$\begin{bmatrix} y_1 \\ y_2 \\ \mathbf{M} \\ y_n \end{bmatrix} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \Lambda & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \Lambda & \varphi_{2n} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ \varphi_{n1} & \varphi_{n2} & \Lambda & \varphi_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \mathbf{M} \\ x_n \end{bmatrix} \quad (3.3)$$

如 $n=3$ 时，上式为

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (3.4)$$

或者

$$\begin{aligned} y_1 &= \varphi_{11}x_1 + \varphi_{12}x_2 + \varphi_{13}x_3 \\ y_2 &= \varphi_{21}x_1 + \varphi_{22}x_2 + \varphi_{23}x_3 \\ y_3 &= \varphi_{31}x_1 + \varphi_{32}x_2 + \varphi_{33}x_3 \end{aligned} \quad (3.5)$$

根据式 3.5 可知， \mathbf{Y} 是 \mathbf{X} 的的线性组合，它通过对各波段的像元灰度值加一个比例系数，综合了 \mathbf{X} 各波段的信息，实现了线性变换，从而使生成的新影像更好地突出实际地表物质的特征^{[34][35]}。

$\sum y$ 是对角矩阵，即协方差矩阵； λ_i (i 为自然数) 是 $\sum x$ 的特征值，即新生成 \mathbf{Y} 的分量方差。

$$\sum_y = \begin{bmatrix} \lambda_1 & 0 & \Lambda & 0 \\ 0 & \lambda_2 & \Lambda & 0 \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ 0 & 0 & \Lambda & \lambda_n \end{bmatrix} \quad (3.6)$$

从式 3.6 可以得知， \mathbf{Y} 的各个分量之间是互不相关的，即新生成的多光谱图像的各个主分量之间是互不相关相互独立的。

根据以上理论，我们对研究区的 TM 影像的七个波段进行 PCA 变换，七个多光谱波段变换成七个主分量，PCA 变换的结果如图 3.1 所示，并且各分量的特征值和贡献率如表 3.1 所示。

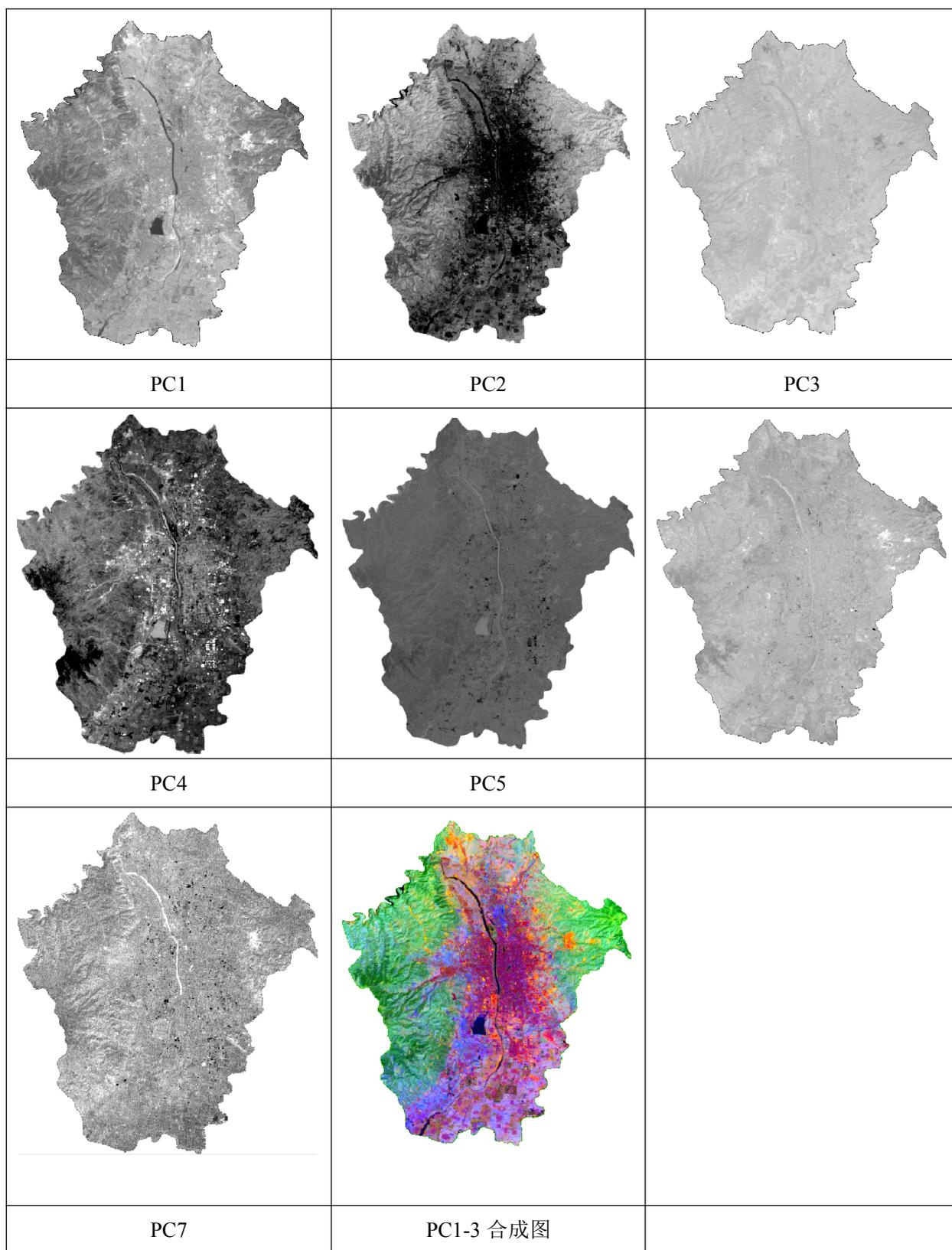


图 3.1 主成份变换结果图

表 3.1 主成份分量的特征值和贡献率

主成分	1	2	3	4	5	6	7
特征值	1235.04	108.53	44.30	20.11	8.61	4.80	0.60
贡献率 (%)	83.74	11.87	2.46	1.38	0.37	0.12	0.06
累计贡献率 (%)	83.74	95.61	98.07	99.45	99.82	99.94	100

从表 3.1 可以看出，变换后的七个主分量分别包含的信息量相差很大，而且表现出不断降低的现象。其中，PC1 包含了绝大部分的信息，占 80% 以上。PC2、PC3 的信息量快速的减少，前三个主成份包含超过 98% 的信息量。到了第 n 分量，信息几乎为零。从图 3.1 我们也可以直观的看出，前三个主分量包含了主要的信息，而剩下的分量几乎为噪声。因为主成份分析对不相关的噪声没有作用，当信息减少时，噪声突出，最后的分量几本是噪声。因此，本文取前三个主分量进行假彩色合成，实现了数据压缩，同时将其作为分类前的特征选择。

3.1.2 缨帽变换

缨帽变换也被称为 K-T 变换，在 1976 年 R.J.Kauth 和 G.S.Thomas 两位学者提出的一种经验性的多波段图像线性正交变换，因而又叫作 K-T 变换。该变换的基本理念：多波段 (N 波段) 可以被视为 N 维空间，每一个像元都是 N 维空间中的一个点，像元在各个波段上的数值决定了它的位置。其变换公式如下：

$$y = Bx \quad (3.7)$$

式中：x——原始多光谱空间的像元值；

y——新坐标空间的像元值；

A——变换矩阵。

缨帽变换旋转后的坐标轴指向与地面景物密切相关的方向，它抓住了实际地表物质

特征，尤其是植被和土壤在多光谱空间的特点。通过研究，在新分量中的前三个分量（图 3.2）与地表物质密切相关，其分别是土壤亮度（SBI）、绿度（GVI）、湿度（WI）。亮度分量（SBI）是 TM 七个波段分量的加权和，表示了整体的亮度变化；绿度分量（GVI）是近红外与可见光波段的比值，垂直于亮度分量，突出可见光波段（尤其是红光波段）与近红外波段之间的比较；湿度分量（WI）与土壤湿度相关，显示了可见光与近红外波段及中红外 5、7 波段的差值，然而中红外 5、7 波段对土壤以及植被的湿度最为敏感。

应用以上理论，对研究区 TM 影像做缨帽变换，获得七维向量空间，前三个分量合成图如下。



图 3.2 缨帽变换结果图

3.1.3 归一化差异水体指数

为了在影像中突出水体，增大水体与其它地物的差异，使用归一化差异水体指数（NDWI, Normalized Difference Water Index）进行了水体提取。

1996 年，Mcfeerers 提出了归一化差异水体指数，其表达式为

$$NDWI = (Green - NIR) / (Green + NIR) \quad (3.8)$$

式中：Green 为绿光波段的灰度值，NIR 为近红外波段的灰度值，对应 TM 影像，分别为第二波段和第四波段。

但是, Mcfeeter 在创建 NDWI 指数时, 仅仅考虑了植被因素, 却疏忽了人工建造物的影响。得到的 NDWI 结果图中, 人工建造物也会具有很高的灰度, 因此, 使用改进归一化差异水体指数 MNDWI^[36], 即 Modified NDWI 指数来提取水体, 其表达式为

$$MNDWI = (Green - MIR) / (Green + MIR) \quad (3.9)$$

式中: MIR 为中红外波段的灰度值, 对应 TM 影像的第五波段, MNDWI 结果在-1 到 1 之间, 结果如图 3.2 所示。从结果图中, 我们可以看出水体得到明显的凸显, 水体信息得到很好的增强。

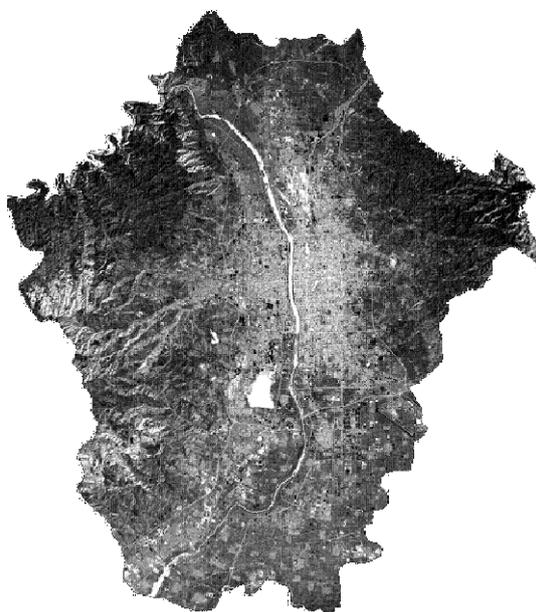


图 3.3 MNDWI 影像

3.1.4 归一化差异植被指数

归一化差异植被指数 (NDVI, Normalized Difference Vegetation Index) 对绿色植被反映敏感, 是植被覆盖率和植被生长状况的指示因子, 被广泛应用于研究中^{[37][38]}。NDVI 之所以被广泛应用于植被遥感中, 有以下三个方面的原因: 第一, NDVI 能够最好的反映植被的生长状况, 表明植被的覆盖情况; 第二, NDVI 经过比值处理后, 可以去除一部分相关的辐照度条件变化等的影响, 如卫星观测角、太阳高度角、云/阴影、地形及大气条件等; 第三, 一些地表覆盖存在明显的特征差异, 水系在近红外波段的反射率低于

可见光波段的反射率，所以它们的 NDVI 值小于 0；而裸地和岩石在这两个波段的反射率几近相等，所以它们的 NDVI 值几近为零；但是对于植被而言，其在 NDVI 近红外波段的反射率高于可见光波段的反射率，所以 NDVI 大于 0，并且植被覆盖度越高，NDVI 的值越大。

在大型覆盖范围下，NDVI 可以很准确地反映光合作用强度、植被绿度以及植被生长的季节和年际变化，因此，该指数被广泛应用于植被的分类以及监测。

根据 NDVI 的指数特征分析，发现利用其可以较为容易地把植被提取出来，其表达式如下：

$$NDVI = (NIR - Red) / (NIR + Red) \quad (3.10)$$

式中：Red 表示红光波段的反射率，NIR 表示近红外波段的反射率，对应 TM 影像的第三波段和第四波段。应用软件建立模型得到 NDVI 结果如图 3.3 所示，

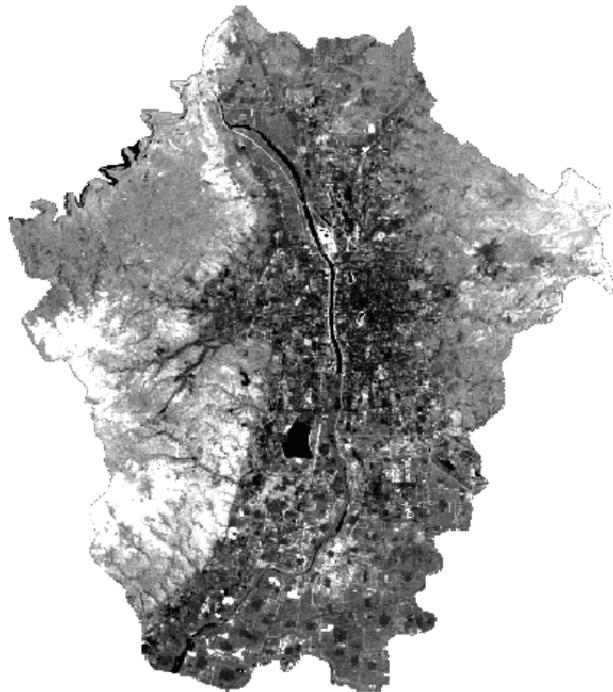


图 3.3 NDVI 指数结果图

3.1.5 归一化差异建筑指数

归一化差异建筑指数（NDBI, Normalized Difference Built-up Index）是查勇等提出的，它是以杨山^[39]提出的仿归一化植被指数为基础。通过研究表面，建筑密度越高，

NDBI 的值就越大，因此，它能够更精确地反映建设用地信息，从而有效地凸显了城市用地信息。计算公式如下：

$$NDBI = (MIR - NIR) / (MIR + NIR) \quad (3.11)$$

式中：MIR 表示中红外波段的反射率，对应 TM 影像的第五波段。通过建模得到 NDBI 结果图 3.4，从图中可以得知，研究区的 NDBI 范围为[-1.0, 1.0]。

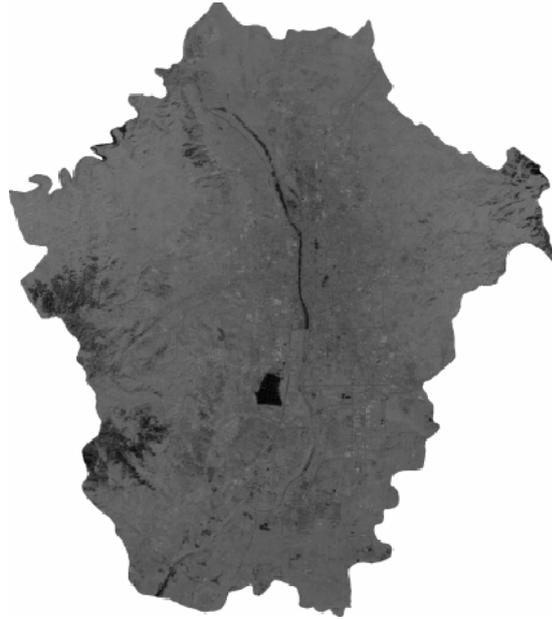


图 3.4 NDBI 指数结果图

3.2 采样点的特征分析

由于地物特征信息的提取是以其光谱特征为基础的，因此在遥感影像上，根据所需分类地物，测量各种地物的每一个波段以及每一个所需特征的像元灰度值，然后从中选定一些典型而且具有代表性的样本点，而后按照各类地物在不同波段样本点的均值做出地物波谱特征曲线。

根据研究需要，我们在本研究中将地表地类分为水体、不透水面、草地、针叶植被和阔叶植被五类。在遥感影像上，就所需分类地物，从中选定一些典型且具有代表性的点，测定其各个波段、PCA 变换的前三个主量的光谱亮度值、缨帽变换后的前三个特征量的样本点的灰度值，而后按照各类地物在不同图层样本点的均值做出地物波谱特征曲线，如图 3.6 所示。

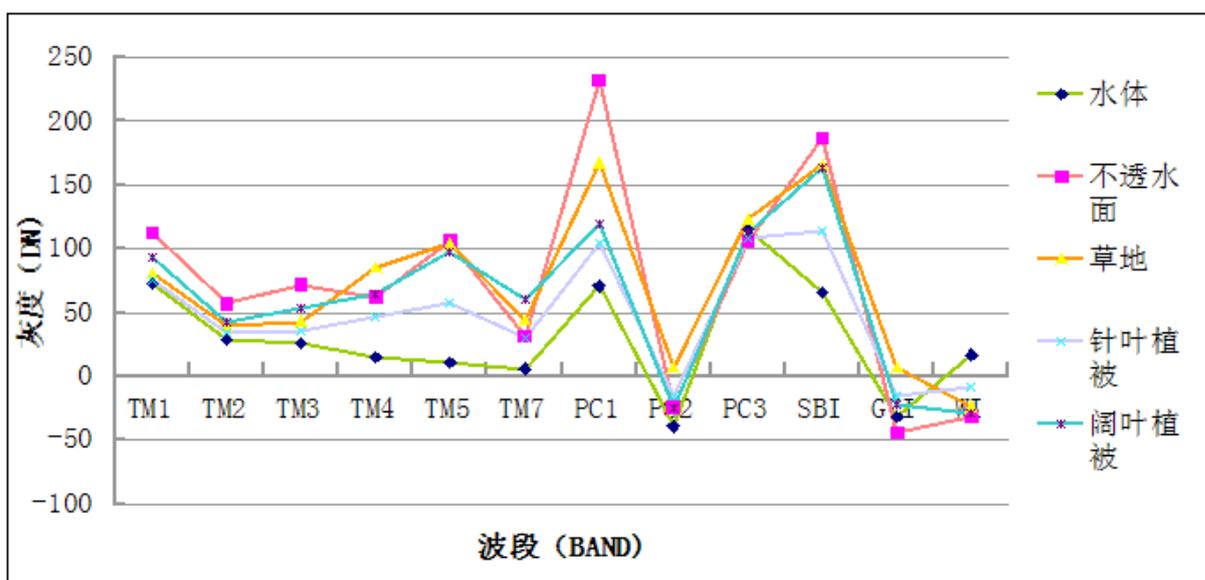


图 3.6 地物波谱特征曲线

根据采样点的光谱特征曲线，我们可以清楚的了解到各种地物的光谱特征差异，进而对其进行分类提取。

3.3 本章小结

本章主要对研究中使用的分类特征进行介绍和分析，讲述了这些特征的原理和获取方法，并通过遥感软件获得这些特征影像。此外，根据特征影像，选取典型且具有代表性的样本点，做出了地物波谱特征曲线，为后续决策树节点的建立做好准备。

4 基于决策树的绿地信息提取

4.1 决策树理论

决策树算法可以处理不同空间尺度的数据,具有计算效率高和无需统计假设等优点,被广泛应用于遥感影像分类领域中^[40]。传统的分类方法分为监督分类与非监督分类。非监督分类方法一般应用于缺乏经验和分类初期,其对遥感数据的分类依据于统计特征,因此分类结果精度较低。监督分类方法首要的是从研究区选取训练样本,然后按照已知训练区域提供的样本,通过选择特征参数,创建判别函数,最后对待分类影像实行分类。监督分类方法假定各个波段中的统计数据满足正态分布,并且对训练样本有较大的依赖性,此外已经建立的提取信息规则不能重复利用。监督分类前必须事先选取训练样本,对于不熟悉地物分布状况的研究者来说,分类前预先选取训练样本存在一定的难度。近21世纪来,随着计算机技术和遥感技术等领域研究的深入,一些新的分类方法相继被提出,决策树是其中之一。

决策树是利用决策规则对遥感数据进行分类的一种数学方法,其利用训练样本进行归纳学习建立决策规则。决策树是一个树型结构,类似于流程图,它由一个根节点、一系列的内部节点和分支与多个叶结点构成,每一个内部节点与一个父节点和两个或多个子节点相连,节点之间经过分支连接。其中,根节点是决策树的开始,位于决策树的顶部,也是决策树的最高层顶点;树的每一个内部节点代表着测试属性,对应一个非类别属性或属性的集合;每一个分支代表不同属性值,对应测试的结果;每一个叶节点就是图像的分类结果,对应一个类别属性值。

决策树方法主要包括决策树学习和决策树分类两个过程。决策树学习过程是以决策树方式产生的分类规则的机器学习过程,建立在对训练样本进行归纳学习的基础上。决策树根据信息论将复杂的决策构成过程抽象成易于理解与表达的规则。其使用信息论中的信息增益找出示例数据库中有着最大信息量的属性字段,建立一条规则,并将此作为决策树的节点。决策树学习算法是由属性和属性值表示的训练样本集作为决策树的输入,规则则作为输出。决策树的建立一般使用自上向下的递归方法,通过一定的方法

选择最优的属性作为树的节点，并比较和判断分支节点的每一个训练样本对应的不同属性的属性值，然后在每个分支重复创建下层结点和分支，且在某种条件下停止树的生长，在决策树的叶节点获得结果，最终生成决策树。根据训练样本进行决策学习建立决策树，然后依据属性的取值对一个未知样本集进行分类，这就是决策树分类^[41]。图 4.1 表示了决策树学习和分类的基本过程框图。

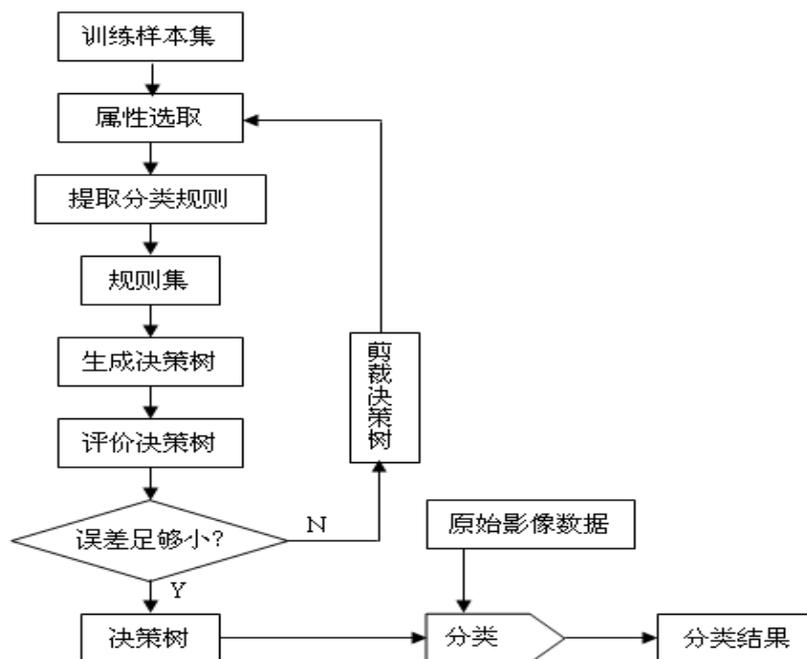


图 4.1 决策树学习和分类的基本过程框图

4.2 绿地提取的决策树模型

应用 ENVI4.6 软件的决策树(Decision Tree)模块，根据采样点的特征分析建立决策树，将地物目标划分为水体、不透水面、草地、针叶植被和阔叶植被五大类。首先我们利用 MNDWI 提取水体，然后利用 NDBI 和 NDVI 提取不透水面，最后根据光谱曲线分析依次提取草地、针叶植被和阔叶植被，以下是利用决策树提取绿地信息流程图（图 4.2）。

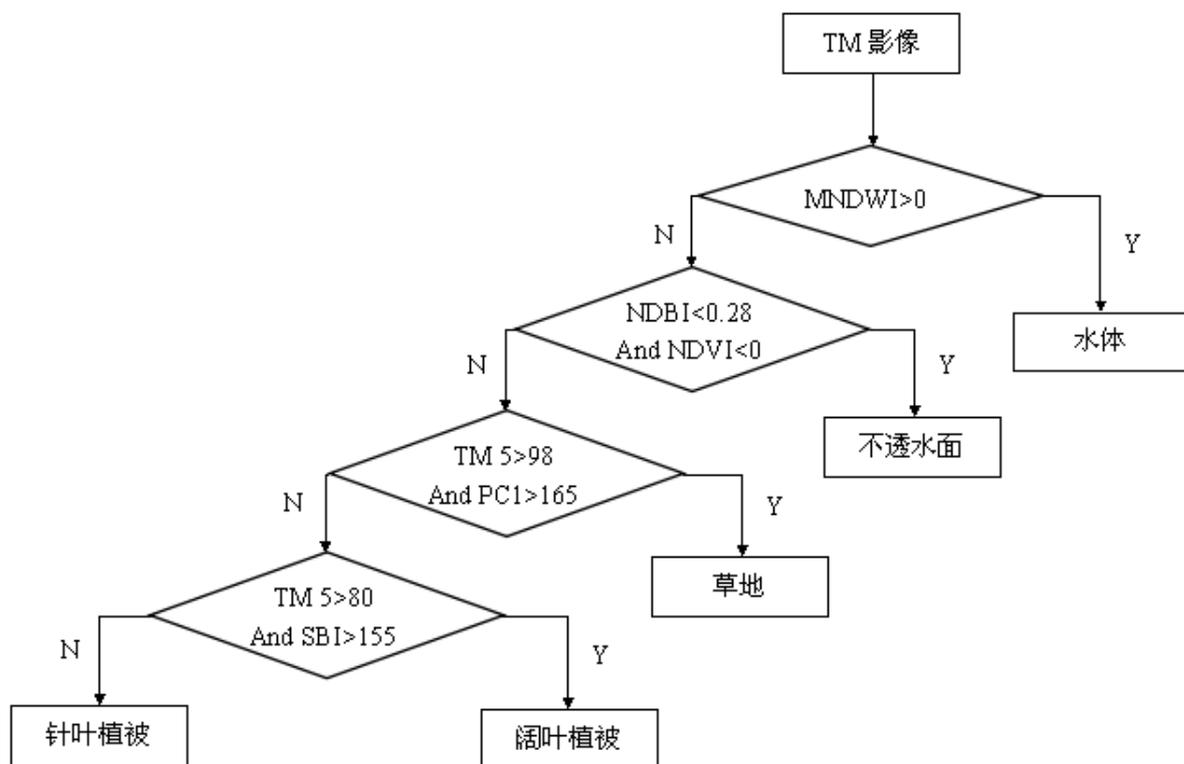


图 4.2 绿地信息提取的决策树流程图

1) 水体的提取

根据 3.1 章的介绍，我们了解到 MNDWI 具有突出水体，抑制其它地物的作用，从 MNDWI 影像图中，我们可以清楚的看到这一点，水体具有很高的亮度值，其它地物相对亮度值低。为此我们首先利用 MNDWI 去提取水体。使用高分辨率影像对比解译，以及根据直方图对比，确定 MNDWI 大于 0 的为水体。



图 4.3 水体提取结果影像

2) 不透水面的提取

去掉水体之后，剩下的地物只有植被和不透水面两大类。我们首先对研究区进行了 NDBI 和 NDVI 的处理，得到影像图。亮度值高的区域是相应指数提取的地表物质，分析 NDBI 影像图，并发现亮色区域中也包含有植被，影响了不透水面信息的提取准确性；而 NDVI 可以有效突出植被覆盖区域，所以结合这两种指数完成对不透水面的提取。同样，使用高分辨率影像对比解译以及直方图对比，确定 NDBI 小于 0.28 并且 NDVI 小于 0 的区域为不透水面。

3) 草地的提取

根据 4.2 章的地物波谱曲线，我们分析发现，草地、针叶植被和阔叶植被同属于植被这一大类，其光谱曲线非常相似，在第一波段到第四波段光谱值基本一致，而草地在第五波段和 PC1 波段较林地（针叶植被和阔叶植被）高出许多，因此可以利用这一特点区分林地和草地。同样，使用高分辨率影像对比解译，确定 TM5 大于 98 并且 PC1 大于 165 的区域为草地。

4) 针叶植被和阔叶植被的提取

根据 4.2 章的地物波谱曲线，进一步分析发现，针叶植被和阔叶植被在第五波段和 SBI 波段有较大差异，尤其在 SBI 波段，为此，我们利用这一特点对二者进行区分，通过对比高分辨率影像，得出，TM5 大于 80 以及 SBI 大于 155 的区域为针叶植被，剩下

的即为阔叶植被。

根据以上分析，最终通过 ENVI 建立决策树模型如下图 4.4 所示，通过运行 Design tree 模块，经过合并和分类处理，并且归并碎小图斑，进行主/次分析处理后，最终形成的分类结果如图 4.5 所示。

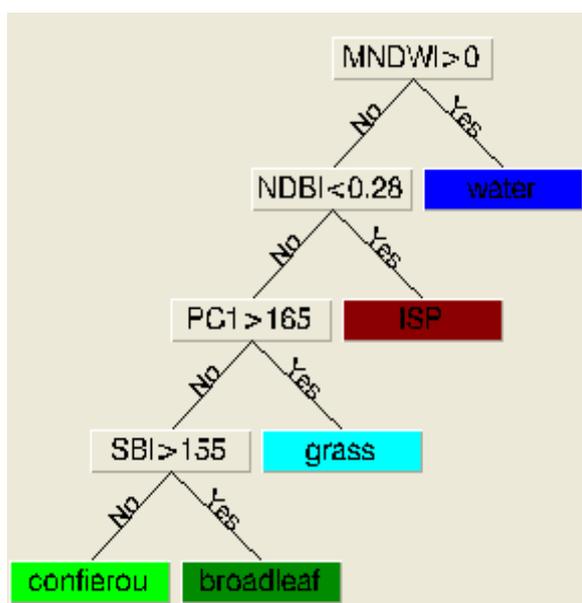


图 4.4 绿地信息提取的决策树模型

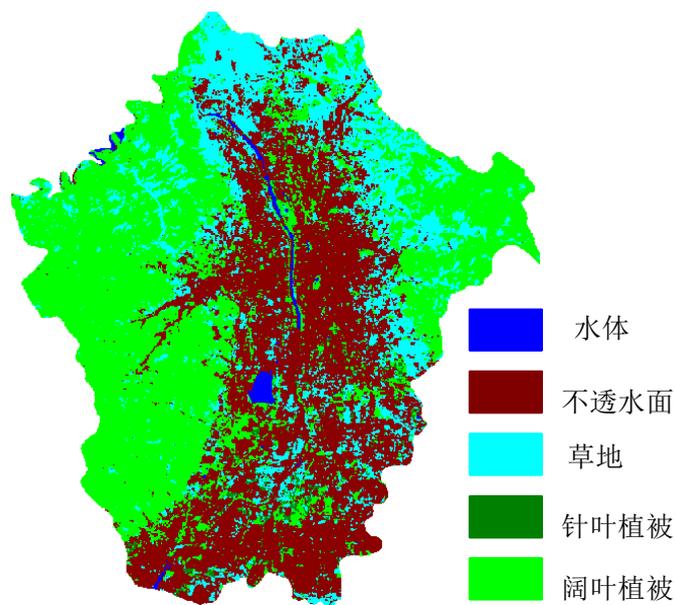


图 4.5 绿地信息提取结果

4.3 提取结果的精度评价

将利用上述方法提取的分类结果通过高分辨率的 ZY3 影像及 Google Earth 影像进行分类精度评估。我们在研究区域范围内随机选择了 200 个样本进行检验,并使用混淆矩阵显示精度验证结果。结果表明,该方法的总体提取的精度可以达到 90%,结果令人满意。

表 4.1 为分类误差矩阵和分类精度评估报告

表 4.1 决策树分类混淆矩阵和分类精度

实际分类 (%)	水体	不透水面	草地	针叶植被	阔叶植被	总体
水体	94.17	0.39	1.93	0.4	3.11	100
不透水面	1.78	92.51	1.85	1.91	1.95	100
草地	1.79	3.22	84.9	4.06	6.03	100
针叶植被	0.22	1.77	6.62	85.25	6.14	100
阔叶植被	2.04	2.11	4.7	8.38	82.77	100
总体	100	100	100	100	100	---

分类总体精度: 85.52% Kappa 系数: 0.837

此外, 我们通过 ENVI 对决策树提取的研究区的分类结果进行面积统计并分析。结果如表 4.2 所示。

表 4.2 决策树分类地物面积统计

地物	水体	不透水面	草地	针叶植被	阔叶植被
面积/km ²	12.72	451.52	300.27	606.12	57.61

应用决策树对研究区域进行遥感分类提取, 取得较好的分类精度。但对于 30m 分辨率遥感数据而言, 由于大量混合像元的存在, 分入绿地丰度高的混合像元的同时, 为保证区域总量精度, 势必会去除绿地丰度较小的像元。这样提取出的结果, 绿地分布是不准确的。因此, 为了进一步提高绿地的面积精度, 需要进行混合像元分解。

为此, 我们将决策树获得绿地总分布图矢量化, 获得绿地分布范围, 做出感兴趣区, 掩膜原始影像, 获得绿地影像 (图 4.6), 进而为混合像元分解减少工作量。

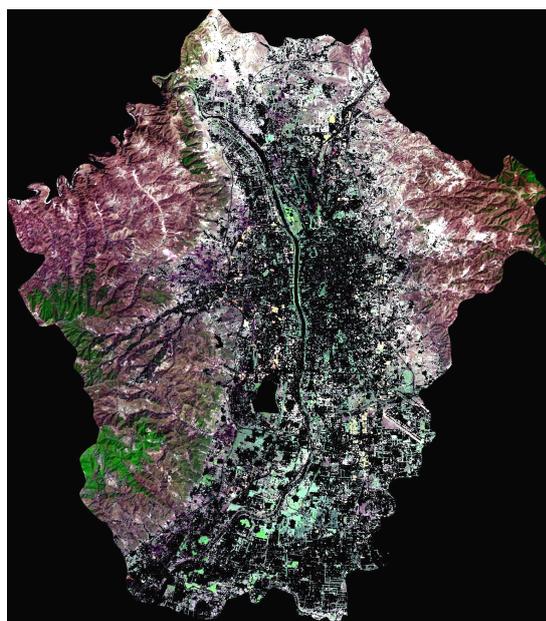


图 4.6 绿地掩膜影像

4.4 本章小结

本章通过分析地物波谱特征曲线，确定节点阈值，建立决策树特征节点，从而构建决策树模型，最终获得地物分类结果图，从而获取绿地分布和绿地面积,以及精度评价。此外，由于混合像元的存在，我们需进一步进行混合像元分解，因此通过绿地分布图矢量化，掩膜获得绿地影像，为混合像元分解提供数据，从而减少计算量和实现高效混合像元分解。

5 混合像元分解

基于决策树的城市绿地信息提取能够较好的提取植被像元，但因为 TM 影像的分辨率为 30 米，使得其面积小于 900 平方米的区域在图像上都是混合像元，这对植被面积的统计测量精度带来很大的影响。

5.1 混合像元分解原理

像元是遥感影像中记录地面的反射或发射光谱信号的载体，它是对应地物的光谱信号的综合结果，其不仅具有一定的光谱参数，而且包含地表物质的空间分布，即相应地物的面积。一般，像元对应的实际地表覆盖物质不是单一的物种，而各个像元只有一个信号用来表示综合地物信号。因此，如果一个像元只包含单一的地表物质，称为纯像元；而如果一个像元包含不止一种地物，则为混合像元。

5.1.1 混合像元分解模型

混合像元的存在阻碍着遥感定量化深入研究的发展，并影响着计算机的处理效果以及其在遥感领域中的发展应用^[42]。然而，传统的像元级遥感分类算法并不考虑混合像元的影响，直接导致传统的遥感分类和面积测量精度不能满足实际需求^[43]。混合像元分解是遥感应用精度得以提高的关键，因此必须进行混合像元分解，从而使遥感应用由像元到达子像元（像元内部），将混合像元分解为不同的端元，并求得这些端元的百分比。

混合像元分解模型一般包括线性模型、几何光学模型、概率模型、模糊分析模型和随机几何模型五种类型，这是由 Charles Ichoku^[44]在 1996 年提出的。线性模型假设，像元的反射率由端元组分的反射率线性组合而得。当忽略多次反射时，线性混合模型是非线性混合模型的特例，归其根本，两者基于相同的概念。对于所有混合像元分解模型而言，端元的光谱特征和混合像元的反射率为必不可少的参数。所有混合像元分解模型最终的分解结果都是各个像元中每一个端元组分的丰度值。线性分解模型是以同一地物具

有相同的光谱特征以及光谱线性可加性为基础，具有结构简单和意义明确的优点，在解决混合像元问题时效果比较理想^[45]，是被广泛运用的混合像元分解模型^[47-49]。

5.1.2 线性光谱混合模型

混合像元是中低尺度影像在城市遥感中普遍感存在的一个问题，简单并且有效得解决这个问题的方法就是线性光谱混合模型，其基于混合像元的光谱是在瞬时视场下的各组分光谱的线性组合，即像元在每个波段的反射率由组成像元的端元反射率以其占像元面积的百分比作为权重系数的一个线性组合。可用以下公式表示：

$$R_{i\lambda} = \sum_{k=1}^n f_{ki} C_{k\lambda} + \varepsilon_{i\lambda}$$

$$\sum_{k=1}^n f_{ki} = 1 \quad (k=1,2,3\dots n) \quad (5.1)$$

式中： $R_{i\lambda}$ 为第 λ 波段第 i 像元的反射率； f_{ki} 为对应于 i 像元的第 k 个端元所占的百分比； $C_{k\lambda}$ 为第 k 个端元在第 λ 波段的反射率； $\varepsilon_{i\lambda}$ 为误差值； n 为端元数目， m 为波段数量，波段数量要大于 n ，使用最小二乘法求解使得误差值最小来获得各类地物的面积丰度值。

线性分解模型具有构模简单、物理意义清晰、理论科学性强等优点，其基于同样的地表物质具有同样的光谱特征，以及像元光谱是地物光谱的线性组合，对于混合像元问题的处理效果比较理想，是目前适用最广泛的混合像元模型。

5.2 混合像元分解模型的建立

本文采用线性光谱混合模型，它是混合像元分解中最成熟、最常用的方法。由于TM数据波段多、信息量大，且相邻波段之间相关性大，存在着信息冗余现象，所以首先进行噪声分离,这样处理后影像噪声降低，数据量减少。其次，端元作为线性混合模型的基本参数，对应具有固定光谱的地表物质，因此利用端元提取来获取图像中的基本信息，在光谱分解、目标检测和图像分类等领域具有重要意义，因此，使用合适的方法进

行端元提取是像元分解的关键。常用的端元提取算法主要有像元纯度指数法、N-FINDR、误差迭代分解和 ORASIS。端元提取算法的选取需根据具体情况选择合适的方法。端元组分的确定即将提取的端元与实际地物进行对应，这是混合像元分解中的难点和重点，它的正确与否在很大程度上直接影像分类精度。通过端元选择技术得到组成光谱后，采用最小二乘法进行计算，最终得到像元各地表覆盖类的丰度。

5.2.1 端元组分的确定

端元选择的途径一般有三种，一种是根据已有波谱库数据进行端元的选择；另一种是从待分解的影像上进行端元的选择；此外，第三种就是二者的结合应用。本研究采用第二种，即直接在待分解影像上进行端元选择。在待分解影像上进行端元选择，除了通过对光谱特征进行目视解译不断尝试外，还可以采用计算机可视化分析工具，即利用散点图和像元纯度指数法进行端元选择。本研究采用像元纯度指数法进行端元组分确定。

1) 最小噪声分离 (MNF) 变换

由于多光谱影像数据的各波段一般高度相关，而混合像元分解的数据源要求相互独立互不相关，因此，需要对影像进行最小噪声分离变换。MNF 变换也被称为噪声白化变换，它是对 PCA 的改进。由两次主成份变换完成 MNF 变换，具体分为三个步骤^[50]：一是对图像进行第一次主成份变换，获得噪声估计协方差矩阵；二是把噪声协方差矩阵转换为单位矩阵；三是对上述处理的图像数据进行第二次主成份变换。

表 5.1 MNF 变换后各波段的相关矩阵

BAND	TM1	TM2	TM3	TM4	TM5	TM6	TM7
TM1	1	0	0	0	0	0	0
TM2	0	1	0	0	0	0	0
TM3	0	0	1	0	0	0	0
TM4	0	0	0	1	0	0	0
TM5	0	0	0	0	1	0	0
TM6	0	0	0	0	0	1	0
TM7	0	0	0	0	0	0	1

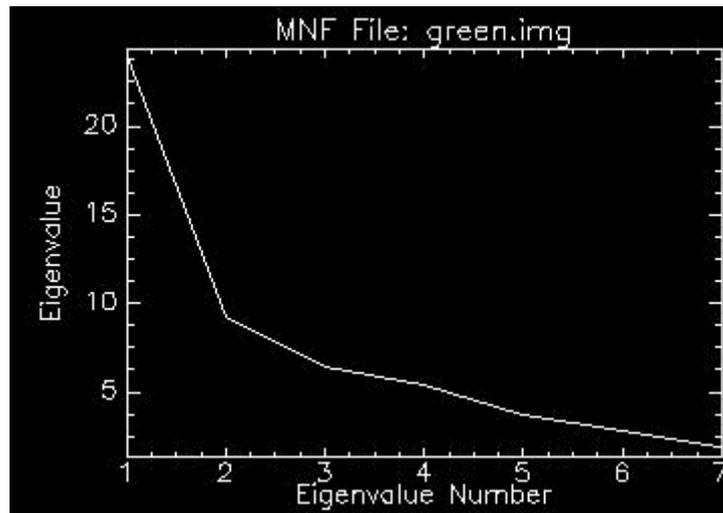


图 5.1 MNF 各波段特征曲线图

从变换后的各波段相关矩阵(表 5.1)可以看出, 经过 MNF 处理, 原始影像的相关性得到了很好的消除, 变换后的 7 个波谱相互独立且各自分布。第一波段的特征值为 23.75, 其方差贡献率为 52.38%; 第二波段特征值为 8.82, 方差贡献率为 22.93%; 第三波段特征值是 6.31, 方差贡献率为 12.41%; 后 4 个波段的特征值分别为 6.19, 3.72, 2.52, 1.19, 方差贡献率分别为 5.79%, 4.71%, 2.32%; 根据这一组统计数据, 可以看出 MNF 变换后的前三个波段集中了原始影像的大部分的有效信息 (81.62%), 而且各种干扰噪声影像从图像中彻底地分离出来, 分成了主成份数据和噪声数据, 原始数据的维数得到了较好的压缩。

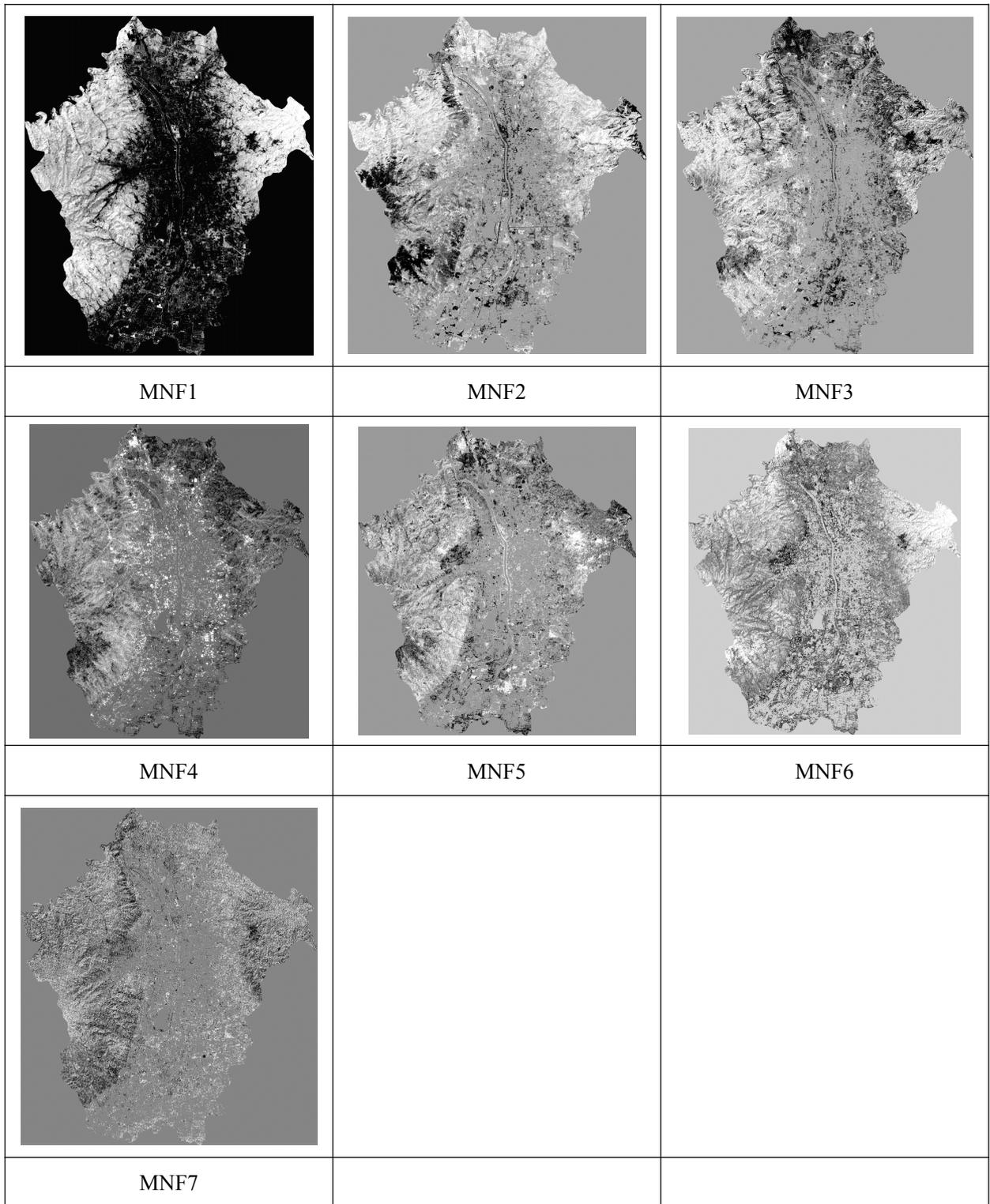


图 5.2 MNF 变换后的各波段分量

从图 5.2 中，我们可以很明显的看出，MNF 变换后的前几个波段几乎包含了原始影像的所有信息，MNF6 和 MNF7 几乎为噪声。

2) 散点图

散点图是形状是不规则的二角洲形、三角形或随机生成的不规则的形状，是向量在灰度空间中的分布，这个向量由不同波段同一像元位置上的 DN 值构成。散点图的顶点对应端元组分，内部点能够用边缘点线性组合得到，边缘点能够由顶点线性组合得到，因此散点图的任意点可由顶点线性组合得到。在进行散点图分析时，为了便于确认端元，应该将原始图像和解译后的 TM 图像交互显示。我们对 MNF 变换的前三个波段两两组合进行散点图分析，分析发现 MNF1 和 MNF2 (图 5.3a) 以及 MNF1 和 MNF3 (图 5.3b) 组合能够较好的获得端元组分，其中，①代表针叶植被，②代表草地，③代表阔叶植被，其余顶点为非植被。但是进一步对二维散点图进行分析发现，但是在进行边缘顶点和 ROI 进行对应分析时，我们发现①、②、③都不是较好的分布在极值点，因此，需要进一步对像元进行提纯。

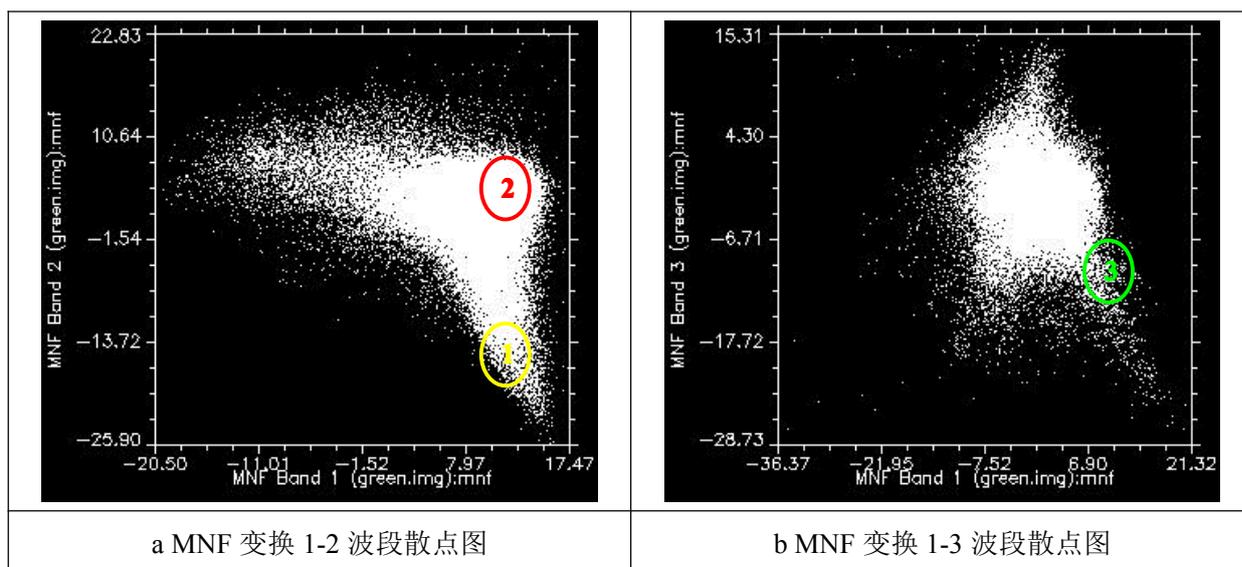


图 5.3 MNF 变换后二维散点图

3) 像元纯度指数提纯

像元纯度指数 (PPI, Pixel Purity Index)用来表示多光谱影像中像元的纯净度。应用 PPI 分析可以在多光谱影像中找出最纯净的像元。PPI 通过迭代将 N 维空间投影为一个随机单位向量，并且记录下各个像元被标记为纯像元的次数。当不断变换随机向量的方向时，记录的次数也不断更新，次数越大，像元越纯净。根据记录次数形成 PPI 指数图，可知，记录的次数对应像元的亮度值，相应地，亮度值越大，像元越纯净。此外，像元

值较高的像元比像元值较低的像元更纯净，所以阈值的合理选择可以提取到更为纯净的像元。

根据以上分析，数据投影到随机向量空间的迭代次数越多，越能够获得更为纯净的像元，但是无疑增加了运行时间。本文设置迭代次数为 1 万次，对 MNF 前五个波段进行 PPI 提纯。之所以只对前五个波段处理，是因为经过 MNF 变换后，最后两个波段几乎都是噪声。PPI 提纯结果如图 5.5 所示。

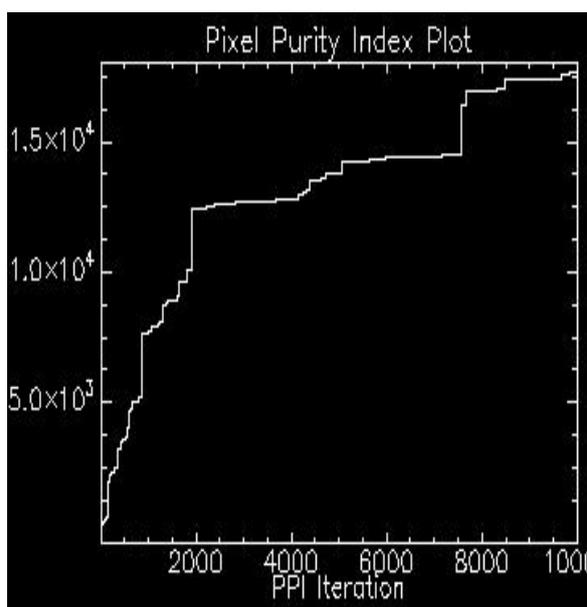


图 5.4 像元纯度曲线图

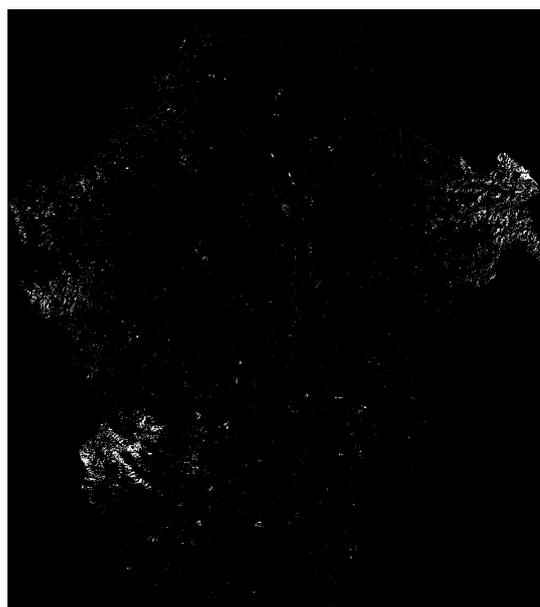


图 5.5 像元纯度图

纯净像元指数法通过反复迭代图像中的像素点，以便找到图像中较纯的像元作为图像端元组分。PPI 指数图中，最纯净的像元就是最亮的像元，因此通过阈值法去除不纯净的像元，保留 PPI 值较高的像元，并将它们提取出来作为 ROI。PPI 分析是为了获得端元像元范围，因此，为了保证完整性和免除遗漏，应尽量选择小的阈值。

4) 获取端元波谱曲线

经过 PPI 提纯后，进一步利用 N 维散点图来纯化端元组分。N 维散点图中的点代表纯像元在各个波段的波谱反射率。研究中将经过 MNF 变换和 PPI 提纯获取的纯像元输入到 N 维空间中，并且不断进行旋转，当类别之间无重叠时停止，获取聚集的像元作为影像的波谱组分。获取的波谱组分输出作为 ROI，从而实现在 N 维空间的端元分离。我们将聚集在一起的区域归为一类，从而获取 N 维散点图（图 5.6）。



图 5.6 N 维散点图

混合像元分解模型的关键就是得到各个地表物质的光谱值，研究中从 N 维散点图中选取点作为端元反射率的值。下面通过 ENVI 利用 PPI 提取的纯端元获取端元波普曲线（图 5.7），为下面混合模型的建立提供基础。

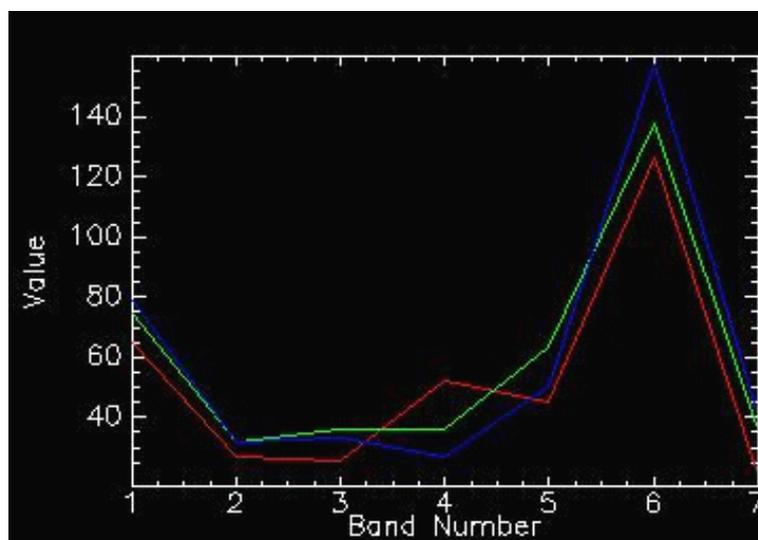


图 5.7 端元波谱曲线

5.2.2 线性混合像元分解

根据波谱曲线获得端元像元值，建立线性光谱模型，利用最小二乘法确定各端元丰度，获得草地、针叶植被和阔叶植被丰度图，其分别的丰度结果如图 5.8。根据获得的

丰度值，再乘以单个像元的面积（900 平方米），最后获得绿地面积。

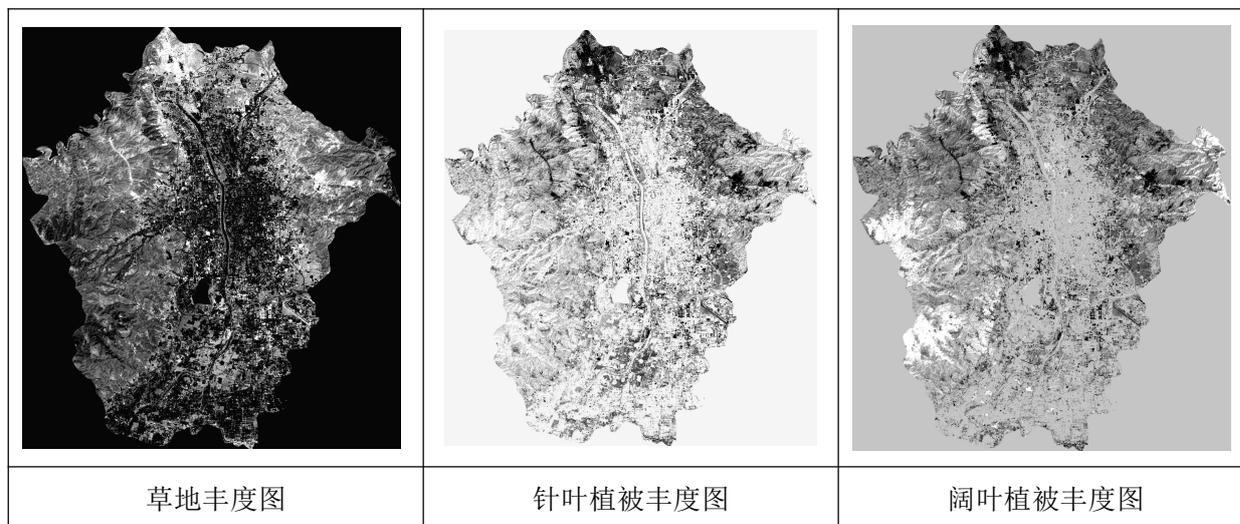


图 5.8 混合像元分解后植被分布丰度图

表 5.2 混合像元分解后绿地面积

地物	草地	针叶植被	阔叶植被
面积/km ²	331.235	392.585	213.742

5.3 精度评价

将混合像元分解的结果通过同期高分辨率的 ZY3 影像及 Google Earth 影像进行分类精度评估，使用混淆矩阵显示精度验证结果。表 5.3 为分类误差矩阵和分类精度评价报告。

为了能够客观的评价决策树和混合像元分解结合的结果，需要一个参照评估处理结果。TM 影像的分辨率为 30 米，为此需要一个更高空间分辨率的影像作为参照，本文选取 ZY3 影像作为参照，其分辨率为 6m。在此以 ZY3 数据监督分类获得的绿地数据作为实际地表情况，从而作为标准数据来评估 TM 数据的分解结果精度。因此，研究中应用监督分类对 ZY3 影像进行分类，将研究区地物分为不透水面、水体、草地、针叶植被和阔叶植被五种地物类型。然后利用此结果评价 TM 影像的绿地面积精度，得出决策树以及决策树和混和像元结合获得的绿地面积结果精度并对比（表 5.3）。

表 5.3 绿地面积混合像元分解前后精度对比表

方法	草地	针叶植被	阔叶植被
决策树精度 (%)	84.90	85.25	82.77
决策树和混合像元 结合精度 (%)	88.95	90.12	87.25

结果表明，经过混合像元分解的绿地面积精度较高，总体精度在 87%以上。经过混合像元分解后，与决策树提取结果精度相比，草地提高 4.05%，针叶植被提高 4.87%，阔叶植被提高 4.48%。可见决策树和混合像元分解相结合提取绿地面积的方法具有较高的精度。

6 总结与展望

6.1 论文总结

本文以 TM 影像作为数据源，结合决策树和混合像元对山西省太原市进行研究，获取城市绿地信息，并统计其面积。我们选择 2011 年 4 月 19 日的 TM 影像和 2013 年 4 月 10 日的 ZY3 影像为主要数据，辅助数据为 1: 250000 基础地理信息数据和实际存储资料作为。首先对获取的遥感影像进行预处理，通过决策树进行绿地信息提取，并进一步利用混合像元丰度解混确定绿地面积。主要工作有以下三个方面：

1) 对 Landsat TM 遥感影像数据进行了预处理。首先对其进行几何校正和剪切，然后为了跟好的获得地物信息，对研究影像进行辐射定标和反射率反演。

2) 利用决策树提取绿地信息。综合分析了主成分分析法（PCA）、缨帽变换、归一化差异水体指数法（MNDWI）、归一化差异建筑指数法（NDBI）和归一化差异植被指数法（NDVI）。综合原图像光谱特征，首先利用水体指数将水体掩膜去除，接着将建筑指数和植被指数相结合去除不透水面，然后对其进行主成分和缨帽变换，分析其波谱特征，最终建立决策树，提取出绿地信息。

3) 对绿地信息进行混合像元分解。首先通过 MNF 和 PPI 变换进行端元提纯，然后通过散点图进行端元组分确定，最终通过最小二乘法进行丰度解混，确定端元丰度。

6.2 不足与展望

1) 决策树特征节点的选取是一个值得探讨的问题，这里我们只用到指数、主成份变换和缨帽变换，在以后的研究中，可以将纹理、高程等特征引入研究中，以及进一步优化特征选取。

2) 混合像元分解时，端元提取是关键，进一步优化端元提取方法是关键。

3) 我们创建的绿地提取方法只是在太原市进行了实验验证，没有进一步在其它城市进行验证，因此，应该在其它城市进行验证，将该方法得到进一步推广应用。

参考文献

- [1] 黄慧萍,吴炳方,李苗苗等.高分辨率影像城市绿地快速提取技术与应用[J].遥感学报,2004,8(1):68-74.
- [2] 陈春林,刘继生,韩阳等.基于TM影像的哈尔滨老城区城市绿地信息提取研究[J].产业与科技论坛,2010,9(5):85-88.
- [3] 王桥,杨一鹏,黄家柱.环境遥感[M].北京,科学出版社,2005.
- [4] Lunetta R S, Johnson D M, Lyon J G, et al. Impacts of imagery temporal frequency on land-cover change detection monitoring[J]. Remote Sensing of Environment,2004,89(4): 444-454
- [5] Thomas H.C Lo et al. Use of Multi temporal Spectral Profiles in Agricultural Land-Cover Classification[J].1986,52(4): 535-544.
- [6] 张治英,徐德忠,孙志东等.遥感图像非监督分类分析江宁县江滩钉螺孳生地植被特征[J].中华流行病学杂志,2003,24(4):259-262.
- [7] Labovitz M.L. Issues arising from Sampling designs and band selection in discriminating ground reference attributes using remotely sensed data[J].PE&RS,1986,52(2):201-211.
- [8] 何国金,刘喜珍等.北京市玉米农情卫星遥感动态监测[J].遥感信息应用技术,2003, 3:28-30.
- [9] Jordan C F. Derivation of leaf area index from quality of light on the forest floor[J]. Ecology, 1969, 50:663-666.
- [10] Kauth R J, Thomas G S. 1976. The tasseled cap-a graphic description of the spectral-temporal development of agriculture crops as seen by Landsat[A]. Pros Symposium on Machine Processing of Remotely Sensed Data[C]. Purdue University,West Lafayette, Indiana: 41-51.
- [11] Deering, D.W. Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. Ph.D. Diss. 1978,Texas A&M Univ., College Station, 338p.
- [12] Huete A R. 1988. A soil-adjusted vegetation index (SAVI) [J]. Remote Sens. Environ, 25:

295-309.

- [13] Baret F, Guyot G, Major D J. TSAVI: A vegetation index which minimize soil brightness effects on LAI and APAR estimation. [C]. Proceedings of the 12th Canadian Symposium on Remote sensing and IGARSS'89 Vancouver, Canada, 1989.
- [14] Qi J A. 1994. Modified soil adjusted vegetation index [J]. Remote Sens. Environ, 48: 119-126.
- [15] Liu H Q, Huete A R. 1995. A feedback based modification of the NDVI to minimize canopy background and atmosphere noise[J]. IEEE Trans Geosci Remote Sensing, 33:457-465.
- [16] FREMBOLD,F.MASELLI.Estimating inter-annual crop area variation using multi-resolution satellite sensor images[J].REMOTE SENSING,2004,25(13):2641-2647.
- [17] 卫亚星,王莉雯,刘闯.基于 MODIS NDVI 时序数据的青海省草地分级[J].资源科学,2008,30(5):688-693.
- [18] 陈燕丽,龙步菊,潘学标等. MODIS NDVI 和 AVHRR NDVI 对草原植被变化监测差异 [J].遥感学报,2011,15(4):838-845.
- [19] Friedl M A,Brodley C E. Decision tree classification of land cover from remotely sensed data[J].Remote Sensing of Environment,1997,61(3):399-409.
- [20] Belward A S,De Hoyos A. A comparison of supervised maximum likelihood and decision tree classification for crop cover estimation from multitemporal LANDSAT MSS data [J]. International Journal of Remote Sensing,1987,8(2): 229-235.
- [21] Schneider A,McIver D K,Friedl M A,et al. Classification of urban areas at continental scales using remotely sensed data[J]. IEEE Geosciences and Remote Sensing Symposium, 2001, 5: 2146-2148.
- [22] Wardlow B D,Egbert S L. Large-area crop mapping using time-series MODIS 250m NDVI data:An assessment for the U.S. Central Great Plains[J].Remote Sensing of Environment, 2008,112(3): 1096-1116.
- [23] 何祺胜,塔西普拉提特依拜,丁建丽.基于决策树方法的干旱区盐渍地信息提取[J].资源科学,2006,28(6): 134-140.

- [24] 孙艳玲,杨小唤,王新生等.基于决策树和 MODIS 数据的土地利用分类[J].资源科学,2007,29(5):169-174.
- [25] 刘永怀.混合像元分解的理论与方法——以神木县为例[J].遥感技术与应用,1992,7(1):7-15.
- [26] Lu D S,Moran E,Batistella M.Linear mixture model applied to Amazonian vegetation classification[J]. Remote sensing of Environment,2003,87:456-469.
- [27] 岳文泽,徐丽华,徐建华等.城市多光谱遥感像元分解技术改进研究[J].浙江大学学报:工学版,2006,40(4):719-723.
- [28] 王学平.遥感图像几何校正原理及效果分析[J].计算机应用与软件,2008,25(9):102-105.
- [29] 姜红梅,任立良,安如等.基于土地利用与地表覆盖遥感信息的洪水过程模拟[J].河海大学学报(自然科学版),2004,32(2):131-135.
- [30] 黄彦丽,蔡楠,李全乐.遥感影像几何精校正实施流程与关键技术探讨[J].科技资讯.2011, 20:60-62
- [31] 党安荣,王晓栋等. ERDAS IMAGINE 遥感图像处理方法[M]. 北京:清华大学出版社.2003.
- [32] 欧春江.实例分析遥感图像处理中的主成分分析[J].测绘与空间地理信息,2006,29(5):56-59.
- [33] 贾海峰,刘雪峰等.环境遥感原理与应用[M].北京:清华大学出版社.2006.
- [34] 陈志强,陈健飞.基于多波段 K-L 变换法的福州市城市用地及其变化信息提取[J].2006,22(3):100-103.
- [35] 秦鹏,陈健飞. K-L 变换与 NDBI 指数法提取 ASTER 影像城市用地信息的比较[J].测绘与空间地理信息,2008,31(5):33-36.
- [36] 徐涵秋.利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J].遥感学报,2005,9(5):589-595.
- [37] 张月丛,赵志强,李双成,孟宪锋.基于 SPOT NDVI 的华北北部地表植被覆盖变化趋势[J].地理研究,2008,27(4):745-754.
- [38] 基于多时相 NDVI 及特征波段的作物分类研究[J].遥感技术与应用,2008,23(5):

520-524.

- [39] 杨山.发达地区城乡聚落形态的信息提取与分析研究—以无锡市为例[J].地理学报,2000,55(6): 671-678.
- [40] Friedl M A,Brodley C E. Decision tree classification of land cover from remotely sensed data[J]. Remote Sensing of Environment, 1997,61(3): 399-409.
- [41] 李亚飞,刘高焕,黄翀.基于决策树分类的云南省迪庆地区景观类型研究[J].资源科学,2011,33(2):328-334.
- [42] 胡争光, 王依婷, 池天河等. 基于混合像元分解和双边界提取的湖泊面积变化监测[J].遥感信息, 2007, 3: 34-38.
- [43] 徐希孺,周莲芳,朱晓红.混合像元的因子分析方法及其在大范围冬小麦播种面积计算中的应用探讨[J]. 科学通报, 1989, 34 (12): 46.
- [44] Charles Ichoku, Arnon Karnieli. A review of mixture modeling techniques for subpixel land cover estimation[J]. Remote Sensing Reviews, 1996, 13: 161 - 186.
- [45] 李剑萍, 郑有飞. 气象卫星混合像元分解研究综述[J]. 中国农业气象, 2000, 21(2): 44~47.
- [46] 延昊, 张国平. 像元分解法提取积雪边界线[J]. 山地学报, 2004, 22(1): 110.
- [47] 岳文泽, 徐建华, 武佳卫等. 基于线性光谱分析的城市旧城改造空间格局遥感研究—以 1997-200 年上海中心城区为例[J]. 科学通报, 2006, 51(8): 966~970.
- [48] 张熙川, 赵英时. 应用线性光谱混合模型快速评价土地退化的方法研究[J]. 中国科学院研究生院学报, 1999, 16(2): 170-172.
- [49] 万军, 蔡云龙. 应用线性光谱分离技术研究喀斯特地区土地覆被变化—以贵州省关岭县为例[J]. 地理研究, 2003, 22(4): 440-443.
- [50] 李素, 李文正, 周建军等.遥感影像混合像元分解中的端元选择方法综述 [J].地理与地理信息科学,2007,23(5):35-38,42.

攻读硕士期间发表的论文

- [1] Wang Feihong, Yang Jiansheng, Li Shiwei, Wang Zhibin. The Research and Application of Medium Resolution Remote Sensing Images in the City Information Extraction , SPIE, 2014, 9142, 91421F1-9.
- [2] 王飞红, 任晓敏. 基于 CBERS-02B 的矿区地物信息的提取[J]. 电子世界, 2013, 121-122.

致 谢

三年的研究生生活马上就要过去了，转眼已到毕业的日子。一路走来，需要感谢的人有很多。

感谢张记龙教授。张老师严谨的治学作风、刻苦的钻研精神、严于律己的态度，深深的激励着我，我将谨记张老师的教诲，在以后的工作和生活中，不断激励自己取得进步。

我要衷心地感谢我的导师杨建生教授。杨老师从论文的选题到最终的完成，倾注了大量的时间和精力，及时解决我遇到的各种难题，使得我的论文得以顺利完成。杨老师严谨的治学态度、缜密的思维方式和平易近人的生活作风，值得我终生学习。

感谢王志斌教授。王老师深厚的理论基础、严谨求实的治学态度、丰富的实践经验、诚恳的为人态度，让我受益非浅。

感谢李晓教授，李老师在学术和生活方面对我的督促，用她独有的干练作风和亲和力教我做人做事。

感谢李世伟老师，不仅在论文的选题到定稿给我提出了很多宝贵的意见，还使我明白许多待人接物与为人处世的道理。感谢杜先荣师妹，感谢光电中心的同门师兄姐妹们，感谢同窗张瑞、王立福、黄艳飞、张鹏飞、杨强、任晓敏、高婷婷，感谢你们在学习和论文写作期间给予的支持和帮助，同时感谢我的室友，感谢她们三年中给予我的帮助。

感谢评审专家对我的论文的评审和指正。

特别感谢这么多年一直给予我支持的父母，没有他们的付出就没有我现在的学业，是他们的关爱和鼓励让我不断进取。

再次衷心地感谢给予我帮助和支持的所有老师、同学、朋友和亲人！