

分类号 TP39

密级

UDC

编号

中国科学院研究生院

硕士学位论文

基于概率模型的高分辨率卫星图像
建筑物识别及变化检测

刘炜

指导教师 普林特 副研究员 博士

中国科学院自动化研究所

申请学位级别 硕士学位 学科专业名称 模式识别与智能系统

论文提交日期 2005 年 5 月 论文答辩日期 2005 年 5 月

培养单位 中国科学院自动化研究所

学位授予单位 中国科学院研究生院

答辩委员会主席 王润生

Building Recognition and Change Inference
Based on Very High Resolution Satellite
Images Using Probabilistic Model

Dissertation Submitted to

Institute of Automation, Chinese Academy of Sciences

in partial fulfillment of the requirements

for the degree of

◀
Master of Engineering

by

Wei Liu

Pattern Recognition and Intelligent System

Dissertation Supervisor: Associate Professor Veronique Prinet

独创性声明

本人声明所成交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确地说明并表示了谢意。

签名：_____ 日 期：_____

关于论文使用授权的说明

本人完全了解中国科学院自动化研究所有关保留、使用学位论文的规定，即：中国科学院自动化研究所有权保留送交论文的复印件，允许论文被查阅和借阅；可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵守此规定）

签名：_____ 导师签名：_____ 日 期：_____

摘要

最近几年，越来越多的商业高分辨率光学卫星图像的出现，给人们带来了新的机遇与挑战。一方面，这些数据能够极大的扩展已有技术的应用范围——比如制图学、城市发展规划与监测；另一方面，人们需要寻找更有效的自动处理方法来提取这些数据中的有用信息。

中分辨率图像的处理主要关注于图像的分类问题，而高分辨率图像的处理则转变为以物体识别为主要任务。对于城市区域的高分辨率卫星图像来说，考虑到分辨率和识别度等因素，道路和建筑物无疑是最重要的人造物体。

目前我们的研究工作是在 863 项目背景下展开的，这个项目的主要目的为快速数字地图自动更新研究开发更有效的技术与方法。本硕士论文主要关注于如下两个问题：单幅高分辨率卫星图像的建筑物识别和两幅高分辨率卫星图像间的建筑物变化检测。本文在对以前遥感图像中物体提取与变化检测技术的分析和总结的基础上，结合项目的应用背景，以北京地区 Quickbird 高分辨率卫星图像为实验数据，围绕建筑物识别与建筑物变化检测进行了研究。本文主要工作及贡献总结如下：

1. 提出了基于概率模型的建筑物轮廓识别算法（第三章）

我们通过定义一个概率密度函数并结合相应的特征来表示建筑物，在执行建筑物识别算法的过程中使用了我们提出的剪切-融合算法。通过不断检测建筑物候选轮廓，衡量它们的概率值，并使用融合算法对某些冲突的轮廓进行合并，以最大化候选轮廓的概率值，从而得到最终的建筑物轮廓。这种算法的优点是能够通过概率模型把建筑物的各个特征进行结合；同时这种算法不仅可以应用于建筑物，也可以应用于其它物体的识别。通过在大数据集上的实验，我们展示了算法的性能及其鲁棒性。

2. 提出了基于轮廓提取的建筑物变化检测算法（第四章）

基于前一章的建筑物轮廓识别算法，我们提出了新的建筑物变化检测算法。这种算法的主要思想是把两幅图像中物体被检测到的概率和没有被检测到的概率的乘积，作为物体的变化概率。这种算法不需要图像的精确配准。我们在合成图像和真实图像上进行实验，展示了算法的效力。

3. 基于 DRF 模型的结构物体变化检测算法（第五章）

为了避免第四章算法中轮廓提取这一主要步骤，我们对已有的 DRF

(Discriminative Random Field) 模型进行了扩展。通过把物体的变化考虑为时间上的不连续性和空间上的连续性,新的 DRF 模型能够直接处理两幅图像间的变化,而不需要首先提取各个图像中的物体。我们在不同图像上进行了试验,展示了这种算法的潜力。

关键词: 建筑物, 物体识别, 变化检测, 概率模型, 高分辨率卫星图像

ABSTRACT

The recent commercial availability of optical very high resolution satellite (VHR) images raises new opportunities and challenges. New opportunities, because these images open the door to a wide range of applications --e.g. cartography, urban planning and control; New challenges, because the automatic processing requires more and more sophisticated techniques to extract relevant information from these data.

While mid-resolution images processing was mainly regarded as a classification task, the problem in high resolution is turned out to an *object recognition task*. For urban applications, among the most accessible --with regard to the resolution-- and easily detectable man-made objects, are the roads and the *buildings*.

The present work was achieved in the context of a 863 project which aimed at providing robust methods and techniques for fast digital map updating in urban area. The specific objective of this Master thesis is to address the *issues of building recognition* from a single VHR image and *building change inference* from a pair of images. This thesis gives at first a general review and analysis of existing methods on object extraction and change retrieval for remote sensing applications (Chapter 2). It then presents in details the newly developed techniques and illustrates results on high resolution Quickbird image covering the area of Beijing. The major achievements and contributions can be summarized as follows:

1. Building contour recognition based on probabilistic model (Chapter 3)

We define the building object by its probability density function --a logistic function-- and its data features. Building recognition is performed using a newly developed cut-and-merge algorithm. This original framework enables to alternately and iteratively detect building contour candidates, estimate their likelihood, then merge possible conflicting candidates such as to maximise their probability to be a building. The advantage of this algorithm is that any data feature can be fused in the model; the proposed method is not specific to the building and can be adapted to any object class. The application on a large data set demonstrates the robustness and performance of the technique.

2. Building change detection based on contour detection (Chapter 4)

We propose a new approach for building change detection based on the building recognition algorithm presented in Chapter 3. The main idea is to compute the change's probability as a product of the probabilities of object

detection and object non-detection calculated in each of the image pair. This approach does not require accurate image registration. We illustrate its efficiency on synthetic and real images.

3. Structural object change inference based on DRF model (Chapter 5)

In order to avoid the preliminary stage of contour detection as required in Chapter 4, we introduce an extension of the existing DRF (Discriminative Random Field) model. The new DRF cope directly with the pair of images and model the change as a discontinuity in time and continuity in space. We performed application and show the potential of the method.

Keywords: building, object recognition, change detection, probabilistic model, high resolution optical image

目 录

第一章	绪论.....	1
1. 1	论文目标及贡献.....	1
1. 2	研究意义与背景.....	1
1. 2	本文组织结构.....	3
1. 3	数据简介.....	4
第二章	遥感图像中建筑物提取与变化检测技术综述.....	7
2. 1	遥感图像中的建筑物提取技术.....	7
2. 1. 1	遥感图像中的物体提取技术.....	7
2. 1. 2	遥感图像中的建筑物提取技术.....	10
2. 2	遥感图像中的建筑物变化检测技术.....	16
2. 2. 1	遥感图像中的变化检测技术.....	16
2. 2. 2	遥感图像中的建筑物变化检测技术.....	20
2. 3	小结.....	24
第三章	基于概率模型的建筑物轮廓识别算法.....	25
3. 1	引言.....	25
3. 2	算法流程.....	25
3. 3	轮廓提取——剪切算法.....	27
3. 4	融合算法及轮廓选择.....	28
3.4.1	轮廓预选择.....	29
3.4.2	融合算法和轮廓选择.....	29
3. 5	模型建立.....	32
3. 6	模型参数优化.....	33
3. 7	特征的选择.....	33
3.7.1	特征图像.....	33
3.7.2	轮廓特征.....	35
3. 8	多边形近似.....	39
3. 9	实验结果及分析.....	40
3. 9. 1	参数设置.....	40
3. 9. 2	结果及分析.....	41
3. 10	小结.....	46
第四章	基于轮廓提取的建筑物变化检测算法.....	49
4. 1	引言.....	49
4. 2	算法流程.....	50
4. 3	配准.....	51
4. 4	模型建立.....	51
4. 5	比较算法.....	52
4. 6	实验结果及分析.....	53

4. 7	小结.....	59
第五章	基于DRF模型的结构物体变化检测算法.....	61
5. 1	引言.....	61
5. 2	算法流程.....	61
5. 3	模型建立.....	61
5. 3. 1	DRF模型回顾.....	61
5. 3. 2	变化检测DRF模型.....	63
5. 4	参数估计.....	64
5. 5	结果推断.....	65
5. 6	特征选择.....	66
5. 7	实验结果及分析.....	67
5. 7. 1	玩具图像结构变化检测试验.....	68
5. 7. 2	合成图像结构变化检测试验.....	69
5. 7. 3	真实图像结构变化检测试验.....	70
5. 8	小结.....	73
第六章	结束语.....	75
6. 1	本文主要工作.....	75
6. 2	未来工作展望.....	76
参考文献		77

第一章 绪论

1.1 论文目标及贡献

本文在对以前遥感图像中物体提取与变化检测技术的分析和总结的基础上,结合项目应用背景,以高分辨率卫星图像数据为基础,围绕建筑物识别与变化检测这一主题进行了研究。本文提出了以下三种算法:基于概率的建筑物轮廓识别算法、基于轮廓提取的建筑物变化检测算法和基于 DRF (Discriminative Random Field) 模型的建筑物变化检测算法。

1.2 研究意义与背景

随着 1957 年 10 月 4 日世界上第一颗人造地球卫星的成功升天,人类开始进入了卫星时代。特别是最近十几年,随着卫星技术的飞速发展,越来越多携带先进技术的卫星不断升空,比如美国的 Landsat 和 Quickbird、法国的 SPOT、中巴合作的 CBERS、日本的 JRS、欧共体的 ERS、加拿大的 Radarsat 等。这些各式各样、功能各异的卫星不断给人们提供着海量的地球观测数据,这就对数据处理工作提出了新的挑战。传统的图像处理技术已经很难应付如此数量的数据;因此,自动、快速而且高效的图像处理技术便成为人们关注的焦点。

在这些卫星图像数据中,高分辨率卫星图像的出现使得在较小的空间尺度上观察地表的细节变化、进行大比例尺遥感制图以及检测人为活动对环境的影响成为可能,因而具有广阔的应用前景。但是,高分辨率图像携带了丰富的信息,如何在这些信息中提取人们关心的信息,是一项富有挑战性的工作,也是图像处理和模式识别领域中的一项重要课题。目前的热点是利用人类对自身环境的认识作为先验知识,对图像中的信息进行提取和分析。

城市是人类活动最密集的地区,也是土地覆盖种类最多、变化最明显的区域。因此,对城市区域的土地覆盖/土地利用进行动态监测是一项非常有意义的工作,可以为资源管理、城市规划和环境变化监测与分析等应用提供有力的决策支持。在城市区域中种类繁多的地物中,建筑物由于分布范围广泛、分布密度高、和人类的生活息息相关,因此成为了城市区域中最重要的人造物体之一。

建筑物的变化检测是城市区域土地覆盖/土地利用动态监测的一项重要工作。

建筑物变化检测，实际上包含着建筑物识别和变化检测两层含义。其中建筑物识别是图像识别的具体应用。所谓图像识别，其实质是个分类过程，即根据遥感图像的光谱特征、空间特征、时相特征，按照解译者或计算机的认识程度，或自信程度和准确度，逐步进行目标的探测、识别和鉴定的过程。这里所谓的“探测”指首先确定一个目标或特征的客观存在，如图像上某处具有一个规则矩形区域等；所谓狭义的“识别”是指在更高一层的认识水平上去理解目标或特征，并把其粗略确认为某个十分普通的、大类别中的一个实体。如确定那个规则矩形区域为人造物体的一部分；所谓“鉴定”是指进一步根据图像上目标的细微特征，以足够的自信度和准确度，经上述“识别”的这个实体，划归到某一特定的类别中，如该人造物体部分是建筑物的屋顶。建筑物的变化检测是遥感图像变化检测的一种特定应用。所谓遥感图像的变化检测，指的是对同一地区，在不同的时间点上获取多幅遥感图像，通过对这些图像进行比较分析，检测出地物的动态变化信息，从而实现对特定区域的动态监测。早期的变化检测主要是通过人工解译，这种方法被公认为是低效率且有遗漏错误的。因此，如何实现自动的非监督的变化检测是当前需要解决的问题。建筑物的变化检测，主要衡量建筑物的三种变化：一是旧建筑物的消失——即在新图像中找不到旧图像中的建筑物；二是新建筑物的出现——即旧图像中不存在建筑物的区域在新图像中出现了新建筑物；三是建筑物形状的改变。

目前，人们往往使用遥感图像和 GIS 结合的方法进行城市区域地物的变化检测工作。其中地理信息系统，即 GIS（Geographical Information System）是计算机支持下管理、处理、分析空间数据的信息系统。对于建筑物的变化检测来说，往往其目的都是为了寻找建筑物的变化以更新 GIS 中相应的建筑物专题图，为 GIS 在其他方面的应用提供准确的数据保障。

对于城市区域建筑物的变化检测来说，高分辨率卫星图像由于能够比较清晰地反映建筑物的基本特征，而且相较于以往的遥感图像获取方法，卫星图像具有更好的经济性和便捷性。因此，在我们的论文中，将选用高分辨率卫星图像作为实验平台。

北京作为祖国的首都，一直都是发展较快的城市，再加上 2008 年奥运会的成功申办，为这座城市注入了新的活力。但是大规模的房屋拆迁，旧城区改造，

城市外扩，道路改造和绿化工程也应运而生。在这样的发展速度下，北京市地图也需要进行及时更新。使用传统的地图更新方法，费时费力，很难满足要求，人们急需一种效率更高的半自动地图更新方法。以中国科学院自动化研究所为主要负责单位，哈尔滨工业大学、中科院遥感所和北京测绘科学院四家科研单位共同承担的国家 863 高科技项目——“城市发展规划中的多元数据融合与决策支持”项目就是在这样的背景下展开的。本文将在该项目背景下，以 Quickbird 高分辨率卫星图像为基础，进行建筑物的提取和变化检测研究工作。在这个项目当中，我们开发了一套能够进行建筑物提取与变化检测系统。本论文所描述的算法也已经实现并嵌入到我们的 863 系统中。

1. 2 本文组织结构

本论文共分六章。

本章主要介绍本学位论文的研究背景与研究意义以及论文的组织结构，并且对论文中的算法将会用到的数据进行简单介绍。

第二章是对当前遥感图像中建筑物的提取和变化检测技术的综述。首先介绍了遥感图像中一般地物的提取方法，然后就建筑物提取技术的发展历史和现状作以分类描述与比较。接下来将会介绍遥感图像中一般地物的变化检测方法，然后具体介绍建筑物的变化检测方法并对一些常见的方法举例说明。

第三章主要介绍我们提出的一种基于概率模型的建筑物轮廓识别方法。在这一章中，将会对这种方法所涉及的主要问题和步骤进行详细说明和分析，最后给出试验结果并作以总结。

第四章主要介绍基于轮廓提取的建筑物变化检测算法。这种新算法是基于上一章的建筑物轮廓识别算法的，因此本章的重点在于描述变化检测概率模型和具体算法。本章最后会给出相应的实验结果并进行分析。

第五章是对我们提出的一个基于 DRF (Discriminative Random Field) 场的结构物体变化检测模型进行描述。这章中介绍的算法还并不成熟，因此我们侧重于对方法的描述和其优缺点的比较。本章会给出算法在不同实验数据集中的实验结果并进行分析，同时对模型的局限性进行讨论并指明今后工作的方向。

第六章是小结。对本文的主要工作和贡献进行了总结，并对未来此领域的工作进行了展望。

1.3 数据简介

本文中使用快鸟（Quickbird）卫星图像作为实验数据。快鸟卫星是由 Ball 航天技术公司、柯达公司和 Fokker 空间公司联合研制，并于 2001 年 10 月 18 日通过美国数字全球公司成功发射的。它是目前世界上空间分辨率最高的商用卫星，其突出的优越性表现在：全色分辨率为 0.6 米左右，多光谱分辨率为 2.44 米左右，是同类卫星 Ikonos 的 1.63 倍；多光谱有红、绿、蓝、近红外四个波段，与 Ikonos 相同；图像幅宽 16.5 千米，是 Ikonos 的 1.5 倍；在没有地面控制点的情况下，地面定位圆误差精度可达 23 米；采用 11 比特/秒数据格式，增加了灰度级数，减少了阴影部分信息的损失。由于以上所示的优越性，目前快鸟图像广泛应用于农作物检测与产量估计、测绘制图、环境监测、军事与国防等多个领域。

在本论文中，我们将会使用北京地区的快鸟图像。由于建筑物识别和变化检测对图像分辨率要求较高，因此我们选择的是全色快鸟图像。文中将会涉及到两个时间的快鸟图像数据，分别获取于 2002 年 3 月 14 日和 2003 年 9 月 13 日。如图 1-1 和图 1-2 所示。在后面的章节中，在没有特别声明的情况下，算法分析和试验中使用的数据都是北京地区 2002 年或 2003 年全色快鸟卫星图像。



图 1-1 2002 年北京地区 Quickbird 全色卫星图像。大小：27552x29014(像素)；分辨率：0.6191 米/像素；成像时间：2002 年 3 月 14 日；左上角地理坐标：(492115, 323842)，右下角地理坐标：(509171, 305881)。



图 1-1 2003 年北京地区 Quickbird 全色卫星图像。大小：27610x30969(像素)；分辨率：0.6264 米/像素；成像时间：2003 年 9 月 13 日；左上角地理坐标：(491236, 323229)，右下角地理坐标：(508531, 303829)。

第二章 遥感图像中建筑物提取与变化检测技术综述

本章主要对当前遥感图像中建筑物提取和变化检测技术进行综述。在 2.1 节中将首先介绍遥感图像中物体提取的基本概念和评价准则，然后针对目前主要的建筑物的提取技术按一定准则进行综合比较和评估。随后的 2.2 节中将首先回顾遥感图像中变化检测的基本概念和技术现状，然后将对目前已有的建筑物变化检测技术作以描述。最后将作以小结。

2.1 遥感图像中的建筑物提取技术

本节首先对遥感图像中的物体识别技术作以回顾。然后将针对遥感图像中建筑物识别技术作以介绍，在这一部分中，将分三个层次对遥感图像中建筑物提取技术的发展作以总结和比较。

2.1.1 遥感图像中的物体提取技术

从图像中进行物体的自动提取是计算机视觉中的一项重要问题。而从遥感图像中提取物体（比如建筑物、道路、水体等）更是一项有意义同时具有挑战性的工作。它是照片解译，地图制作，GIS (Geographical Information System) 数据库更新等众多应用的关键技术。

本章中所谓的“提取”，主要包括“识别”和“重建”两重含义。识别指从图像中通过相对简单的特征和照相机模型找到感兴趣的物体的位置，并使用简单的 2D 或者 3D 模型进行表示；重建则表示利用已知的物体几何知识特别是其拓扑结构，结合高质量的数据和更为复杂的照相机模型对物体进行表示。

根据 Mayer 在[1]中的描述，一般遥感图像物体提取方法可以根据图像数据、模型的选择和策略的使用三个方面进行分类：

I. 图像数据

根据提取物体的目标和应用目的的不同，所采用的图像数据及其复杂度也各不相同。图像数据按种类可大致分为：航空图像和卫星图像。图像数据的复杂度主要包括内容和可观察性两方面：

① 内容

密度：乡村/郊区/城市

分布复杂度：居住区/工业区/军事区

建筑风格：精细/普通

地形：平坦/山地/丘陵

植被：无/稀疏/密集

② 可观察性

分辨率：一般提取不同的物体使用不同的图像分辨率，根据[1]，表 2-1 例示了提取不同物体需要的图像分辨率。一般来说，高分辨率的卫星图像相当于中分辨率的航空图像。

图像质量：对比度/噪声强度/季节因素

同一场景的不同图像：单幅/立体/多幅/彩色

对于不同的地物，可以利用光谱信息的加以区别。多幅图像为物体三维重建提供了信息，特别是对于建筑物。

物体种类	分辨率
建筑物	2m
小径	2m
支路	5m
水体	5m
主路	10m
街区	10m

表 2-1 提取不同物体需要的图像分辨率

II. 模型

- 典型特征一般反映物体的功能[3]。这些典型特征把三维真实世界中的知识整合到模型中，将会对提取物体最终的语义特性起到重要作用。

- 建立材料特性模型[4]用于反映传感器特性，减少其负面影响。当有不同传感器时，这点尤其重要，比如同时具有光学、雷达传感器。
- 二维几何和拓扑特性[5]。特别适用于道路和建筑物提取。对于建筑物提取，它们更适合于简单形状的建筑物，比如长方体形的建筑物，而对于更复杂形状的建筑物，这种模型并不完善。
- 详细的图像特征模型[6, 7]。利用图像属性特征生成的连接图，可以把图像中的特征联系起来。这种模型尤其适用于多幅图像的匹配，相对于仅仅利用单一特征，能够提高匹配正确率。
- 抽象级别和尺度[3]。这种模型主要用于道路提取和植被提取。也可能适用于建筑物提取技术。
- 几何的临域相关/拓扑的临域相关[4]。就是指物体空间的上下文关系。比如汽车与道路相关，影子和建筑物联结等。
- 各部分结构[8, 3]。也称为子结构。比如建筑物上的窗户，门，屋顶的烟囱，天窗等。
- 统计模型[9]。考虑数据的不确定性，利用概率统计理论提取物体。目前这方面比较热门，最新的技术倾向于使用贝叶斯网络和马尔科夫随机场理论。

III. 策略

- 基于外貌的方法[10]。这种方法避免了从图像模型到物体模型的转换。比如对于树木，很难用物体模型来建模。当图像分辨率特别高时，这种方法可以用来提取屋顶的细节。
- 分组[5, 11]。也就是说，在搜索几何和拓扑规律时，可以关注于物体的子结构，从而减小搜索空间。一般来说，分组应和验证配合使用。
- 使用多尺度[3]。可以先在较低分辨率对物体进行提取，找到物体大概位置和轮廓，然后在较高分辨率上进行细化。这种方法对物体提取具有明显的加速作用。
- 利用其他已探测物体的上下文关系，产生假设物体的预测，然后通过搜索和再分割的方法验证是否真正存在此物体[12]。这种方法适用于那些较难提取和需要语义信息的物体。

- 从子结构中获取证据[10, 3, 6]。一般的子结构很难被直接提取，但是当物体被提取后，其子结构可以在物体区域的限制下加以提取，这样提取的子结构可以反过来对提取的物体加以验证。比如对于建筑物的门窗等细致结构。
- 数据融合和算法融合[13, 7]。数据融合指对彩色，多光谱等来自不同传感器的数据进行融合。算法融合是一个广泛的概念，比如对于多尺度空间或者边界和区域的不同定义方式等。

2. 1. 2 遥感图像中的建筑物提取技术

城市区域的遥感图像往往包含大量复杂而且高密度的物体，比如建筑物，街道，公园等人造物体和植被，水体等自然景物。在这些物体中，建筑物一般被认为是最突出的特征物体，归因于其数量，复杂性和在人类社会活动中的重要意义。因此，城市区域图像的建筑物提取一直是人们关注的焦点。

Herman 和 Kanade 的方法[12]是早期经常被引用的方法，他们使用基于人工智能的 3D 推断，搜索垂直和水平方向的直线，运用启发式推理提取长方体建筑物。一些年后，建筑物模型依然使用长方体假设，但是，出现了通过对垂直或者平行边界分组获得的矩形进行匹配，产生 3D 结构的方法[11]。Shufelt 的方法[5]可以从单幅，倾斜视角的图像中提取平顶或者尖顶的建筑物，他主要使用影子信息和基于消逝点的垂直边界。Wang[6]证明建筑物的语义信息不仅仅局限于几何范围。从校正过的航空图像上提取墙面的窗户和门等精细结构，可以提高建筑物提取的正确率。Weidner[13]展示了如何利用数字表面模型（DSM: Digital Surface Model）提取类似于尖形，棱柱形的参数化建筑物。Fischer 和 Henricsson[8, 7]分别提出了基于细致的图像模型和复杂策略的，对多幅图像的底层特征，比如边界和拐角进行匹配的方法。Fischer[8]通过把建筑物的各部分进行参数化表示，构造了一个一般性的建筑物模型，与此对比，Henricsson[7]并不使用一个完整的模型，而仅仅试图找到部分的屋顶。Krishnamachari 的方法[14]首先提取图像中的直线段，然后基于直线段构造 Markov 场反映直线段之间的关系，通过最小化能量函数找到矩形和“L”型的建筑物轮廓。Gerke[15]使用图像光谱特性找到建筑物大概区域，然后利用直方图的方法分离出建筑物和其阴影区域，最后使用几何不变矩重建 2D 建筑物轮廓形状，他的方法可以提取复杂多边形轮廓的建筑物。Garcin 和 Descombes[16]提出了超越 Markov 场的

Marked points 场并结合 RJMCMC (Reverse-Jump Markov Chain Monte Carlo) 进行建筑物提取, 这种方法的好处是能够把建筑物的几何特征和关系集成到能量函数中去。Stassopoulou[17]使用 Bayesian 网络的方法进行建筑物提取, 其优点在于能够综合考虑建筑物几何、纹理、照相机模型和场景模型及其之间的关系, 并且采用先验知识进行学习。Lin 和 Nevatia[18]构造了一个非常典型的从底层到顶层的建筑物提取方法, 使用图像中的边缘、墙面和阴影信息结合假设验证的方法提取矩形立方体建筑物。Haithcoat[19]采用 LIDAR 数据生成 DSM 然后基于建筑物的高度特性并结合几何特征进行建筑物的提取。Wei 和 Zhao[20]的在单幅图像上进行建筑物提取工作, 首先利用灰度信息进行聚类, 然后使用建筑物的基本几何特征并结合阴影进行假设验证, 最后使用 Hough 变换重建 2D 建筑物轮廓。Tison 和 Tupin[21]的方法比较特别, 他们使用的数据是高分辨率的 SAR (Synthetic Aperture Radar) 图像, 通过 SAR 图像本身的特性, 构造阴影和建筑物间的能量函数, 反映它们之间的关系。Guo 和 Sawhney[22]结合模型驱动和数据驱动学习在同一场景多幅航空图像上进行建筑物提取, 首先建筑物的大概位置和轮廓被检测出来, 然后利用通过先验知识学习得到的树型分类器对初始建筑物提取结果进行验证和优化。

我们沿用上一节的分类方法, 依然从三个方面来论述遥感图像中建筑物提取技术的发展历史和现状。我们主要使用上面介绍的具有代表性的建筑物提取方法进行分析。

I. 图像数据

表 2-2 显示了上述各种方法的数据类型和数据复杂度——主要分为分辨率和图像内容两个方面。早期由于技术的原因, 主要处理航空图像, 随着时间的发展, 更多类型的数据出现了, 特别是高分辨率卫星图像的出现, 为建筑物提取技术的发展增加了更广阔的空间。目前的技术发展趋势是在高分辨率高复杂度图像上进行建筑物提取。

II. 模型

除了上一节中描述的适合于一般遥感图像中物体提取的模型, 针对遥感图像中建筑物提取的模型有如下几种形式:

- 影子和墙[5]对于单幅图像来说是很好的房屋 3D 表示依据。但是也存在着很

多问题：影子投影到不平的地面，被建筑物本身遮挡，投影到其他建筑物；在垂直角度墙面信息很少，被别的建筑物遮挡等。

- 多幅图像中的 3D 几何模型[8, 7]。利用照相机模型，可以找出图像中的 3D 物体。
- 通用的 3D 房屋模型[8]：由表面模型和 CSG (constructive solid geometry) 模型组成。

表 2-3 展示了 16 种建筑物提取方法在模型使用上的不同。主要考察几何/辐射特性、传感器模型的复杂度、建筑物模型、场景模型和建筑物子结构的复杂度几个方面。从表 2-3 可以看出，大部分算法都主要使用几何特征，只有少数算法使用辐射特性；在传感器模型的选择上，一般需要进行 3D 重建的算法会使用较为复杂的传感器模型，这是由于 3D 重建需要进行投影变换，对传感器模型参数的要求也较高。对于航空图像，由于其具有较高的分辨率，能够反映建筑物的细部特征，因此在建筑物模型的选择上也倾向于选择复杂的 3D 建筑物模型，并经常考虑建筑物的屋顶形状；但是对于卫星图像，由于一般高分辨率的卫星图像仅相当于中分辨率的航空图像，所以在卫星图像上进行建筑物提取一般只考虑 2D 建筑物模型。在场景模型的选择上，很多算法都应用了阴影，这是由于图像中阴影是建筑物高度特征的最直接的体现；同时我们可以看到，Stassopoulou[17]的算法考虑了复杂场景，其中包括建筑物周围物体和建筑物之间的位置关系，这种复杂关系是通过 Bayesian 网络模型进行描述的，这也是建筑物提取算法未来发展的方向之一。在子结构的选择上，大部分算法都不考虑建筑物的子结构，这是由两方面原因造成的：某些算法所使用的数据分辨率的限制；算法本身无法描述建筑物和其子结构件复杂的拓扑关系。

III. 策略

除了上一节中描述的适合于一般遥感图像中物体提取的模型，针对遥感图像中建筑物提取的策略有如下几种形式：

- 在多幅图像中融合底层特征[8, 7]。通过匹配不同角度图像中的底层特征，可以更好的进行三维重建。而且，应该尽量利用 DSM 这一有力工具。
- 平面形状的使用。也就是说，首先利用平面形状构造物体空间模型，然后：3D 点=>房屋局部=>房屋=>匹配图像底层特征和局部房屋模型。

- 从 DSM 中提取的房屋假设[13]。这种策略并不是很可靠，因为 DSM 提取的假设往往并不正确，比如由于植被的影响，因此只能作为进一步分析的一个依据。

表 2-4 描述了各种算法中策略的使用情况。主要从融合、分组、控制/搜索方式和评价几个方面作以总结和比较。可以看出，由于航空图像本身的特性，一般都倾向于使用统一场景的多幅图像进行建筑物提取，同时一些图像则结合了 DSM 等特征图像；而卫星图像则多使用光谱特性，有些算法直接使用了雷达图像。分组算法仍然是大部分算法中的一个重要的部分。而在控制/搜索方面，早期多采用数据驱动算法，而现在的倾向是使用模型驱动算法，这是由于建立模型能更好的集成先验知识，如果再结合学习算法，则能让先验知识决定模型参数，这种途径显然更符合人类认知的规律。在算法结果的评价方面，大部分算法都采用相对简单的评价方法，这是由于一直以来都没有一个针对建筑物提取的评价准则。直到 2005 年，Song[2]提出了一套相对较为完备的建筑物提取评价体系，在本人的工作中，将会采用这套评价体系。

方法	数据类型	数据复杂度	分辨率	图像内容
Herman and Kannade[12] 1984	航空	低-中	中	简单
Mohan and Nevatia[11] 1989	航空	中	中	中
Shufelt[5] 1996	航空	中	中	中
Krishnamachari[9] 1996	航空	中	中	中
Wang et al. [6] 1997	航空	中	中	中
Weidner[13] 1997	航空	中	中	简单-中
Fischer et al. [8] 1998	航空	高	高	中-复杂
Lin and Nevatia[18] 1998	航空	高	高	中-复杂
Henricsson[7] 1998	航空	高	高	中-复杂
Stassopoulou[17] 2000	航空	高	高	中-复杂
Gerke[15] 2001	卫星	高	高	中-复杂
Garcin and Descombes[16] 2001	航空	高	高	中-复杂
Guo and Sawhney[22] 2001	航空	中	高	中-复杂
Haithcoat[19] 2001	卫星	高	高	中-复杂

Wei and Zhao[20] 2004	卫星	高	高	中-复杂
Tison and Tupin[21] 2004	卫星	中	高	简单

表 2-2 数据复杂度比较

方法	几何和 辐射	传感器 模型	建筑物模型	场景模型	子结构
Herman and Kannade[12] 1984	几何	简化	简单(棱柱)	无	无
Mohan and Nevatia[11] 1989	几何	简化	简单(棱柱)	无	无
Shufelt[5] 1996	几何	详细	中(平顶+尖顶+ 墙面)	简单 (阴影)	无
Krishnamachari[9] 1996	几何+辐射	详细	中(3D 模块+ 子 结构)	简单 (阴影)	简单; 隐含 (门, 窗户)
Wang et al. [6] 1997	几何	-	简单(2D 矩形)	无	无
Weidner[13] 1997	几何	-	中(标准建筑物 类型+复杂平顶)	无	无
Fischer et al. [8] 1998	几何	详细	详细(复杂 3D 结 构, 平面形状)	无	无
Lin and Nevatia[18] 1998	几何	详细	详细(复杂 3D 结 构)	无	无
Henricsson[7] 1998	几何	简化	中(3D 模块+ 子 结构)	简单 (阴影)	无
Stassopoulou[17] 2000	几何	详细	中(2D 多边形)	复杂(周 围环境)	无
Gerke[15] 2001	几何+辐射	-	简单(2D 矩形)	无	无
Garcin and Descombes[16] 2001	几何	-	详细(复杂 3D 结 构)	无	无
Guo and Sawhney[22] 2001	几何	-	中(2D 多边形)	无	无
Haithcoat[19] 2001	几何	-	中(标准建筑物)	无	无

			类型+复杂平顶)		
Wei and Zhao[20] 2004	几何+辐射	简化	中(2D 多边形)	无	无
Tison and Tupin[21] 2004	SAR	详细	中(2D 多边形)	阴影	无

表 2-3 模型的使用

方法	融合	分组	控制/搜索	评价
Herman and Kannade[12] 1984	不同时间图像	3D, 简单	数据驱动/-	无
Mohan and Nevatia[11] 1989	立体	3D, 复杂	数据驱动/ 约束满足	简单
Shufelt[5] 1996	-	3D, 复杂	数据驱动/-	简单
Krishnamachari[9] 1996	-	2D, 复杂	模型驱动/-	无
Wang et al.[6] 1997	多幅图像	-	模型驱动/-	无
Weidner[13] 1997	-	-	数据驱动/-	复杂; MDL
Fischer et al.[8] 1998	多幅图像	3D, 复杂	假设验证/ 约束满足	复杂
Lin and Nevatia[18] 1998	单幅图像/ 多幅图像	3D, 复杂	假设验证/ 约束满足	简单
Henricsson[7] 1998	多幅彩色图像, DSM	2D + 3D; 复杂	数据驱动/-	简单
Stassopoulou[17] 2000	-	2D, 复杂	模型驱动/-	无
Gerke[15] 2001	-	-	数据驱动/-	无
Garcin and Descombes[16] 2001	多幅图像	-	模型驱动/-	简单
Guo and Sawhney[22] 2001	多幅图像	-	模型驱动, 假设验证/-	简单
Haithcoat[19] 2001	立体/DSM	-	数据驱动/-	复杂
Wei and Zhao[20] 2004	-	-	数据驱动/ 约束满足	无
Tison and Tupin[21] 2004	-	-	模型驱动/-	无

表 2-4 策略的使用

2. 2 遥感图像中的建筑物变化检测技术

本节首先对遥感图像中一般物体的变化检测方法作以分类和回顾，然后就已有的建筑物的变化检测方法作以简单描述。

2. 2. 1 遥感图像中的变化检测技术

在本文中，遥感图像变化检测问题的具体含义是：在两幅或多幅单波段或多波段覆盖地面同一区域，并在不同的时间（相隔几年或几个月）获取的遥感图像上，找到变化的区域和没有变化的区域并分析和确认变化区域的变化类型（例如原来的田地后来成为居民区等）。变化主要有两种含义：一是物理上的变化——即真实世界的变化；二是图像反映的变化——包括真实世界的变化在图像上的反映，也包括图像本身由于其它因素造成的变化，比如噪声、拍摄角度、阴影等造成的图像变化，这些变化并不是真实世界的变化。变化检测的核心问题是如何排除其它因素对图像变化的影响，从而确定两幅图像中特定地物发生的真实变化。在我们综述中的方法中，有一些多时相（至少三时相）多光谱图像的变化检测的方法，但变化检测最基本的概念是两时相图中的对应像像素是否发生变化。因此问题最简单的描述是：对于给定图像 I_1, I_2 ，怎样确定那些明显发生变化的区域或物体。同时，也要尽量去除那些不感兴趣的变化，如由传感器噪声或气候光照的不同带来的伪变化。根据彭文辉的硕士学位论文的总结[23]，遥感图像变化检测方法可以分为两类：

➤ 基于分类的方法

✓ 分类后比较法

首先运用统一的分类体系对多时相遥感数据组中的每一时相遥感图像进行单独的分类，然后根据分类的结果逐个像素比较或是对分割得到的区域进行比较，直接得到变化信息。分类后比较法的关键是高精度的空间配准与准确的分类。其优点是：（1）有利于减少多时相数据因获取季节不同和传感器差异所造成的误差；（2）图像经单独分类后比较，可直接得到变化的类型、数量和位置；（3）分类后比较法是单独分类，无时相数目的限制，可以同时

进行两个时相以上遥感图像的变化检测。其缺点在于：(1) 无法检测出存在于某一土地覆盖类型内部的细微变化；(2) 由于分类的误差可能会夸大变化区域，分类的错误全部被当作莫须有的变化；(3) 二时相分类数据进行比较后生成的变化图，其精度只相当于每个时相数据分类精度值的乘积。

✓ 直接多时相分类法

这种方法是将两个或多时相的遥感图像数据放在一个数据库中同时进行分类。出现变化的类别，其数据的统计量如标准偏差等将很大，而未发生变化的类别其数据的标准方差很小，因此通过统计量的比较可以区分出类别变化与否。可以看到，这是一种比较复杂的方法，因为它常常需要很多类和特征。而这其中有一些很可能是冗余信息，可以通过主成分变换等方法去除冗余。另外一个问题是，在合成的数据库中，时间特征和光谱特征具有同等地位，因此，分类中很难将光谱变化和时间变化轻易区分开来。

➤ 基于光谱特性的方法

这一类方法是通过对不同时相图像像元的光谱差异分析，来判断地物是否发生了变化。其具体的操作过程是：通过诸如图像差值、图像比值、主成分分析（PCA）、植被指数差值等一些不同的数据变换方法，将多时相的原始图像变换成一个新的单波段或多波段的图像，然后对此图像进行分析从而区分变化/未变化像元。对提取出的变化像元可以作进一步的分析，如分类或集群化等来确定其具体的变化类型，也可以直接用高分辨率的图像或其他辅助资料来确定变化类型。

下面，将具体对现有的遥感图像变化检测方法作以分类描述：

✓ 图像差值法[24]

这种方法将已配准的两个时相的遥感图像按波段进行逐像元的相减。差值图像中的数据分布的总体特征是：高亮度的部分一般是发生了强烈变化；低亮度的部分或是归一化不彻底引起的变化。单变量图像差值法自然影像的变化检测来说过于简单，很难考虑到所有因素的影响，同时容易造成信息容量的损失。

✓ 图像比值法[25]

这种方法将已配准的多时相遥感图像按波段进行逐像元的相除，然后逐像元比较。比值法可以部分消除阴影的影响，突出某些地物之间的反差，具有一定的图像增强作用。比值法的缺点在于比值图像的数据通常不是正态分布，划分变化/未变化像元的阈值不好定义；同时可能加强在各波段之间互不相关的随机噪声和相关噪声的影响。

✓ 图像回归法[26]

图像回归法首先假定时相 t_1 图像的像元值是另一个时相 t_2 图像的某一个线性函数，通过最小二乘法来进行回归，然后利用回归方程计算出来的预测值减去时相 t_1 的原始像元值，从而获得二时相的回归残差图像。经过回归处理后的遥感图像数据在一定程度上类似于进行了辐射水准的相对校正，因而能减弱多时相数据中由于大气条件和太阳高度角的不同所带来的不利影响。Mark 在 1997 年 [26] 中对遥感图像变化检测中的回归法进行了分析。

✓ 植被指数差值法[27, 28]

植被指数差值法是用近红外与红光波段间的比值（植被指数）来代替原始波段作为输入数据进行差值运算来生成变化图像。比较常用的植被指数有比值植被指数、归一化植被指数和正交植被指数等。植被指数法广泛应用于土地植被覆盖的变化检测中[28]，可以取得较为精确的检测结果。

✓ 主成分分析法

主成分分析法的数学基础是离散 Karhunen—Loeve 变换（简称 K—L 变换）或霍特林变换（Hotelling），是建立在图像统计特征基础上的多维线性变换。多光谱图像经 K—L 变换后，其主成分可以作为图像差值法的输入参数，即：每个时相的多光谱数据先独立作 K—L 变换，然后取其主成分图像进行比较分析，以求出变化信息。PCA 方法的优点是能够分离信息并减少相关性，从而突出不同的地物目标；不足之处在于变换产生的成分其物理意义有时并不明确。

✓ 典型相关分析法[29]

与 PCA 方法类似，典型相关分析法的思路是，将多时相的多光谱图像看作是多元随机变量，引入多元统计分析的理论和方法进行变化检测的研究。其实质是把差异总信息分配到互不相关的几个变量上，以达到最大限度保持这一差

异的总信息量不改变的情况下，检测出图像之间发生的变化。

✓ 变化向量法[30, 31]

陈晋等在 2001 年[31]提出一种参考图像分类结合变化向量方向余弦最小距离分类的方法。其思路是：在对参考图像的准确分类的基础上，求取变化类型对应的光谱差值向量在方向余弦空间中的特征点，根据各变化像元对应变化向量与特征点的相似性（用距离最小来描述）来判定其变化类型。

✓ 基于 Bayes 决策论的变化检测算法[32, 33]

L. Bruzzone 在 2000 年[33]提出一种基于 Bayes 最小错误率的变化检测方法。首先，两幅已配准的图像直接作差（如果是单波段则像素值直接相减，如果是多光谱图像则作矢量差，取其幅值分量），生成差值图像，然后对差值图像中的每一个像元按变化/未变化分为两类，将像元划分到对应后验概率最大的那一类中去。

✓ 基于图像分割的变化检测算法[34]

L. Bruzzone 在 2000 年[34]提出一种基于区域块的变化检测的方法。首先对二时相图像分别进行分割，每幅图像均分割得到若干个区域，再对分割出来的小区域进行融合，使两幅图像最终获得对应相同的区域块分割。综合考虑图像像素灰度值和纹理信息，定义每一个分割区域的特征向量，基于此特征向量生成二时相图像的差值图像，最后选取合适的阈值来确定变化区域。这种方法的优点在于考虑了图像的空间信息，避免由于去噪处理带来的检测误差。缺点是前期图像分割结果对后期变化检测的精度影响很大。

✓ 基于三维分割的变化检测算法[35]

T. Yamamoto 在 2001 年[35]提出一种应用三维分割来检测多时相多光谱图像变化的方法。将图像坐标系和时间轴结合起来形成一个三维坐标系，这样多时相图像在三维坐标系中形成一个大的图像方块。其前提条件是多时相图像已经完全配准，且经过波段归一化处理（各幅图像具有相同数目的波段，且各波段波长范围相近）。这样，在图像方块中，垂直于时间轴的光谱特征的边界即代表了时间特征上的光谱辐射值的变化，这就将检测变化转换成搜索图像方块的光谱边界的问题上。

✓ 基于小波变换的变化检测算法[36]

L. M. T. Carvalho 等在 2001 年[36]提出一种将多尺度小波分析应用于变化检测的方法。方法是首先通过重采样使两幅图像的像素尺度相同 然后, 由这两幅图像生成差值图像, 再对差值图像采用以三次样条为尺度函数的小波分解算法, 得到四幅新的分量图像, 分别反映原差值图像的概貌和细节信息, 对其进行分析, 选取合适的分量图像作假彩色合成, 可以直接目视判别出变化区域, 避免了阈值的选取。

✓ 基于马尔可夫场的变化检测算法[37]

T. Kasetkasem 等在 2002 年[37]提出一种基于马尔可夫场 (MRF: Markov Random Field) 模型的变化检测的算法。其基本思想是: 给定两幅原始图像, 最终要得到一幅二值变化图像区分变化与非变化, 如果将这一问题看成是假设检验的问题: 每一个假设对应一幅不同的变化图像, 其中满足最大后验概率 (MAP) 的变化图像, 即为最优的检测结果。该方法避免了前人在差值图像上应用马尔可夫模型[33]的做法, 但是这种方法具有相当的计算量。

✓ 基于概率分布的变化检测算法[38]

Jordi Inglada 在 2003 年[38]中提出一种适用于两时相合成孔径雷达灰度 SAR 图像变化检测的方法。他提出, SAR 图像的统计特性可以由 pearson system 综合描述[39], 因此, 判断图像像素在不同的时间是否发生变化, 可以通过判断其统计特性是否发生变化来实现。

✓ 基于假设检验的变化检测算法

Srinivas Andra 等人在 2003 年一篇图像变化检测综述的初稿中提出, 几乎所有变化检测问题都可以看成是统计检验的问题: 判断给定像素在两时刻是否发生变化, 就是在两种假设中做出决策: 一种是原假设或零假设——对应像素没有发生变化, 一种是备选假设——对应像素发生了变化。这种决策涉及到代价损失函数的评估和决策阈值的选取。大多数方法中, 代价损失函数都是基于差值图像来考虑的。

2. 2. 2 遥感图像中的建筑物变化检测技术

在遥感图像变化检测中，建筑物的变化检测是一项十分重要的问题。建筑物的变化检测在城市规划、灾害分析、地图更新等方面发挥着重要的作用。由于建筑物是一种特定的地物，因此针对建筑物，人们也研究开发出了相应的变化检测方法。这一小结将主要介绍建筑物的变化检测方法，通过对几种典型的建筑物变化检测方法进行简单描述，展示这一技术领域的发展和现状。

类似于建筑物提取，在进行建筑物变化检测时，我们主要利用航空图像或者卫星图像。由于航空图像具有较高的分辨率，并且能够提供统一场景的多角度图像，因此在以往的建筑物变化检测算法中，多使用航空图像。随着技术的发展，现在卫星图像也能够提供较高的分辨率，并且具有很好地经济性，因此现在越来越多地建筑物检测算法越来越多地使用卫星图像数据。

真实世界中建筑物的变化在图像上的反映一般具有三种形式：一是旧建筑物的消失——即在新图像中找不到旧图像中的建筑物；二是新建筑物的出现——即旧图像中不存在建筑物的区域在新图像中出现了新建筑物；三是建筑物形状的改变——这可能包括很多种情况：比如建筑物部分结构的消失/出现、建筑物几何形状的改变等。由于对第三种情况的检测需要较详细准确的描述建筑物拓扑结构，这是一个困难的问题，因此目前大部分算法主要关注于前两种变化情况的检测。

利用图像进行建筑物的变化检测，会受到很多因素的影响，这就决定了建筑物的变化检测是一项具有挑战性的工作。一般来说，主要的影响因素有：

- ✓ 光照条件的影响：由于图像来自于不同的时间，因此成像时光照条件的差异会造成图像灰度分布的差异，同时，太阳照射角度的不同也会造成阴影投射方向的变化。
- ✓ 拍摄角度的影响：对于航空图像来说，拍摄角度的差异也会造成物体投射的变化。
- ✓ 季节因素的影响：比如夏冬季的季节差异在植被分布上的反应——夏季相对茂密的树木可能会对建筑物产生遮挡。
- ✓ 人为因素的影响：比如，未发生结构变化的建筑物，也可能由于人为的因素在辐射特性上发生非结构特性的变化——比如屋顶颜色的变化等。

由于建筑物是特定的物体，因此建筑物的变化检测需要完成两方面的工作：一是区分出建筑物和其它地物；二是找到建筑物间的变化。要完成这两项工作，可以采取两种顺序：先找到图像的变化区域，然后在变化区域中寻找建筑物，则找到的这些建筑物即是发生变化的建筑物；或者，先在图像中找到建筑物，再寻找它们之间的变化。目前来说，绝大部分算法都是按照第二种顺序进行的，实际上，先找到建筑物再检测它们之间的变化相当于上一节中介绍的基于分类的变化检测方法。下面，将介绍一些具有代表意义的建筑物变化检测方法。

✓ Watanabe 的方法[40]

Watanabe 和 Miyajima 的方法是应用于航空图像上的。这种方法的核心思想是建立一个 3D 光照模型，寻找 2D 图像的变化。这种方法避免了寻找图像上建筑物的 3D 结构——因为往往这是一个费时费力的过程。他们的光照模型如下所示：

$$I(i) = \rho(i)(L_e + L_d \cos \theta(i)) \quad \text{公式 2-1}$$

其中 $I(i)$ 表示位置 i 的像素强度， L_e 和 L_d 分别表示环境光照和直接光照， $\rho(i)$ 是 Lambertian 反射参数， $\theta(i)$ 是光源和位置 i 像素的法线交角。定义了光照模型后，计算两幅图像像素强度的比值： $I_1(i)/I_2(i)$ ，并且把这个比值与预先根据建筑物和阴影区域反射特性定义的阈值进行比较，这样能够同时提取出发生变化的建筑物区域和阴影区域，通过再次阈值能够区分开建筑物区域和阴影区域。这种方法的关键是阈值的选取，定义好的阈值能够获得好的结果，在他们的方法中并没有实现阈值的自动选择。

✓ Huertas 的方法[41]

不同于传统的方法，Huertas 的方法并不使用两幅不同时相的图像进行比较，他基于 Gerson 和 Wood 提出的 Site Model[42]，首先假设旧图像中的建筑物已被提取成为 Site Model 组成的数据库，新图像与数据库中模型进行比较，找到变化的建筑物。这种算法的主要步骤如下所示：

- ① Site Model 和新图像的配准：使用线段匹配的方法把旧图像的 Site Model 和新图像进行配准。
- ② Site Model 验证：通过简单特征确认 Site Model 在新图像中的位置上是否

真的存在物体。

③ 结构变化检测：使用更详细的特征来验证上面得到的 Site Model 和新图像物体间的对应关系。

④ Site Model 更新：根据新算法确认的物体变化情况，更新 Site Model 数据库，为下一次变化检测做准备。

Huertas 的方法的优点是能够使用先验知识，并且整个系统可以不断更新。缺点是算法只能够找到消失的物体，无法找到新出现的物体。

✓ Agouris 的方法[43]

Agouris 的方法是一种典型的使用 GIS 数据库作为先验知识进行建筑物变化检测的算法。这种算法的关键是使用最小平方距离衡量新图像中建筑物边界和 GIS 矢量多边形的距离。其算法流程图 2-1 如下所示。

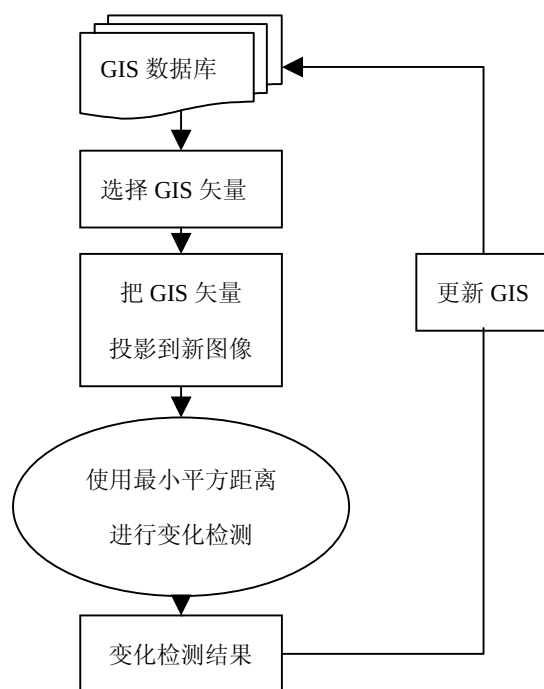


图 2-1 算法流程图

✓ Shintaro 的方法[44]

Shintaro 的方法[44]使用的数据是不同拍摄角度、不同时间的航空图像。通过寻找两幅图像中对应的建筑物屋顶进行变化检测。在他的方法中，主要关注的是由于图像拍摄角度不同带来的视差和建筑物阴影对变化检测的影响，为了消除这种影响，使用了照相机模型。其中照像机模型的参数是通过手工标记对应点儿估算出的。Watanabe 的方法的贡献在于相对以往的多视角变化检测方法，额外考虑了视差和阴影的影响，因而获得了更好的实验结果。

✓ Thuy 的方法[45]

Thuy 使用 LIDAR 数据进行地震地区的建筑物变化检测。针对 LIDAR 数据的特殊性，基于这种数据的建筑物变化检测方法也具有其特殊性。Thuy 的方法主要步骤为：

- ① 预处理：主要是把不同时相的 LIDAR 数据统一到一个尺度范围内
- ② 主处理：Thuy 并不是基于像素点的，而是根据 LIDAR 数据点的分布密度，定义小格子为基本的数据单元。对相应格子中的不同时相的 LIDAR 数据进行相减得到它们的差值，然后基于这些差值生成直方图，那些和直方图均值的差大于某个阈值的小格子便是发生变化的建筑物区域，其中的阈值采用人工指定而没有自动选取。
- ③ 后处理：在主处理中被认为发生了变化的格子中，考察不同时相的 LIDAR 数据点的具体变化情况，对之前的检测进行验证和校正。

这种方法的好处是充分利用了 LIDAR 数据的优点，能够获得较好的结果。不足之处是算法中多处用到的阈值都是人工定义的，没有实现自动选取。

2.3 小结

本章主要对建筑物识别及其变化检测的基本概念作以描述，并对这一领域的发展历史和技术现状作以总结。可以看出，目前还没有任何一种算法能在所有图像数据上都获得较好的结果，人们往往针对具体的应用，开发相应有效的算法。在本文接下来的章节中，将会介绍针对我们的项目背景，应用于特定数据上的建筑物识别算法和变化检测算法。

第三章 基于概率模型的建筑物轮廓识别算法

3.1 引言

本章将介绍基于连续区域最外层闭合轮廓提取并结合概率模型对建筑物进行识别的算法。这种算法的核心是通过建立概率模型来衡量某个闭合轮廓区域是否为建筑物的可能性，并结合我们提出的剪切-融合算法对候选闭合轮廓进行选择。下面将首先展示整个算法的流程，然后将详细描述流程中各个步骤。最后，将会给出试验结果并对其进行分析。

在我们的算法中，寻找的是建筑物的屋顶，以后建筑物轮廓也是指建筑物屋顶轮廓。另外在本论文后面的描述中，如无特殊指明，“轮廓”均代表连续区域的最外层闭合轮廓。

3.2 算法流程

类似于其它基于像素、直线或者区域的物体识别算法，本章介绍的算法是基于“轮廓”这个基本元素的。我们提出了“剪切-融合”算法：“剪切”指对输入图像采用不同阈值进行二值化分割并从中提取出一系列候选轮廓；“融合”指对轮廓进行合并处理。之所以采用“剪切-融合”算法，是因为通过简单的图像阈值，很难直接得到图像中物体（建筑物或非建筑物）的准确轮廓，因此我们希望首先获得一些轮廓，这些轮廓可能不能单独描述一个完整的物体，但是它们可能分别属于某个物体的一部分并且某些轮廓区域之间具有重叠部分。这样，我们可以通过“融合”操作对某些重叠区域进行合并得到正确的物体轮廓区域。

为衡量某个轮廓区域是否为建筑物轮廓区域，我们引入了一个概率模型，这个模型通过计算轮廓区域的某些特征，能够给出轮廓是否为建筑物轮廓的概率值，进而我们可以挑选出那些属于建筑物的轮廓区域。对于概率模型中的未知参数，我们利用先验知识（训练集）结合参数优化方法对其进行优化，从而使模型具有良好的分类能力。

图 3-1 给出了整个算法的流程图。图中“轮廓提取”即“剪切”算法；“轮廓融合”和“轮廓选择”之所以捆绑在一起，是因为在算法实现上它们密不可分；最后的“多边形近似”起到对建筑物轮廓矢量化的作用。

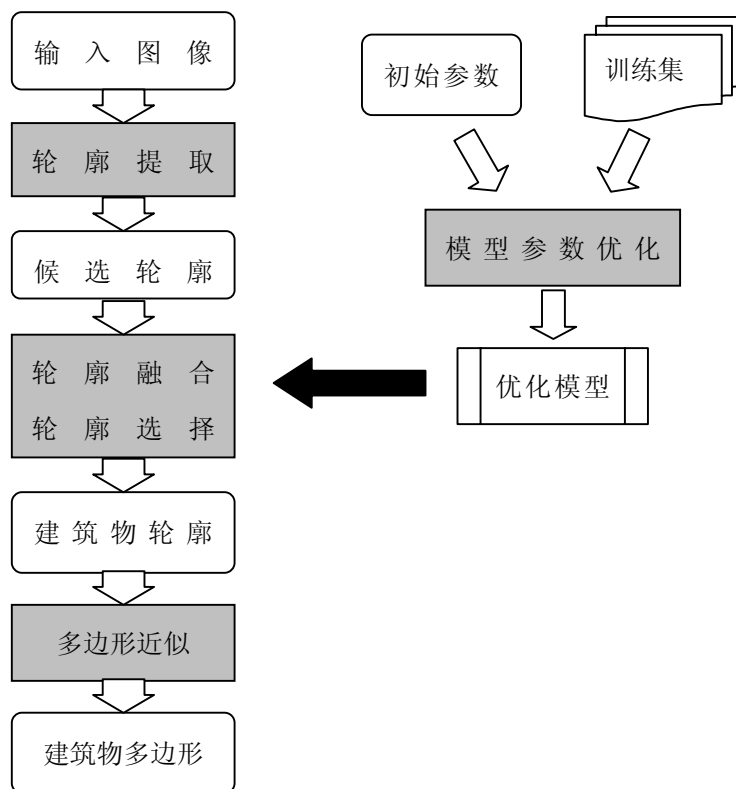


图 3-1 算法流程图

后面章节将介绍算法流程中的各个步骤，以图 3-2 所示图像为输入图像。



图 3-2 原始图像。大小：256X256，分辨率：0.6 米/像素，拍摄时间：2002 年

3.3 轮廓提取——剪切算法

剪切算法的目的是提取图像中的轮廓以构成候选轮廓集，为下一步的处理提供基本“元素”。这里我们采用“多滑动窗口二元带通滤波器”对输入图像进行带通二元滤波，生成一系列二元图像并进行去噪声处理；基于这些二元图像应用轮廓检测算法，找到图像中各个连续像素区域的外层轮廓，所有这些轮廓构成整个候选轮廓集： $C = \{C_i\}_{i \in S}$ ，其中 S 是轮廓空间。

所谓“滑动窗口二元带通滤波器”，即预先定义滑动窗口大小，按照一定的步长，从限定的灰度级最小值滑动滤波器窗口直至限定的灰度级最大值。每滑动一次即在原始图像上应用一次带通滤波器，只有处于窗口灰度级范围内的像素会被保留下来，构成一幅二元滤波图像。所谓“多滑动窗口二元带通滤波器”即定义多个滑动窗口大小。经过滤波后的二元图像像素区域不够连续，存在很多细小的孤立区域和连续区域中的小缺口（黑白噪声），我们分别使用数学形态学中的“开”、“闭”算子去掉黑白噪声。

剪切算法相应的伪代码如下所示：

```

procedure CUT_PROCESS(image, slid_win_len[], slid_win_num, step,
                        min_level, max_level)
for i = 1 to slid_win_num
    for j = min_level to max_level - slid_win_len[i] by step
        bin_image = up_down_threshold(image, j, j+slid_win_len[i])
        bin_image = morphologic(bin_image)
         $C = C \cup \text{find\_contours}(\text{bin\_image})$ 
    repeat
repeat
end CUT_PROCESS
  
```

在此算法中，控制各个输入参数可以得到不同的区域候选集。由于建筑物本身的材质多种多样，反射率各不相同，因此在图像中所表现出来的灰度级分布也比较广泛，对于各个参数的设定应该在保证算法效率的同时尽量宽松，以保证能够获得尽量多的区域以覆盖所有建筑物区域。同时我们也可以看出，经过剪切算法得到的区域候选集中存在着大量具有重叠区域的区域，这也是我们

后续采用融合算法的原因之一。图 3-3 是剪切算法的一个实例。

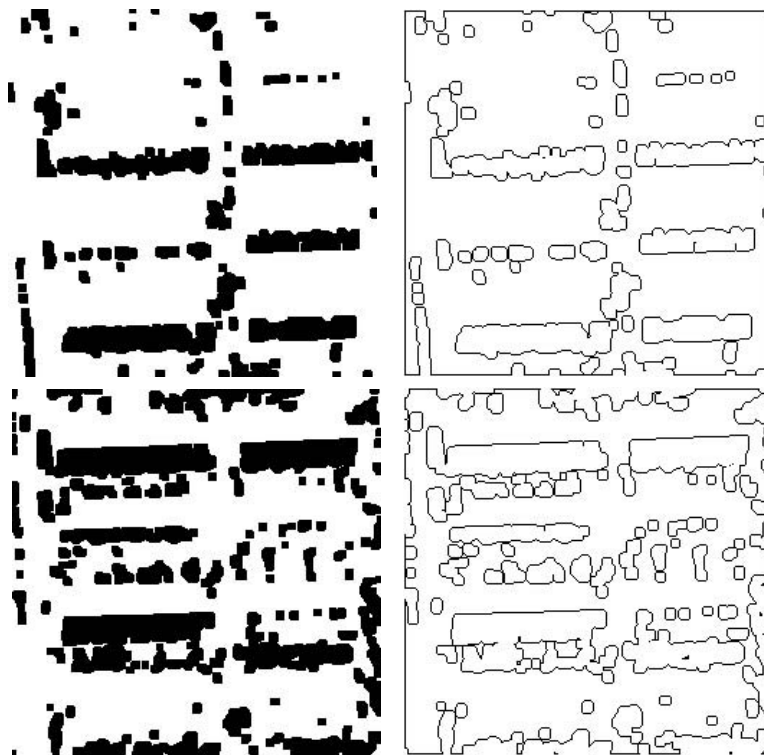


图 3-3 在图 3-2 上应用剪切算法的结果。左上为窗口宽度 30，在灰度级范围 110 至 140 时的滤波结果，右上是相应的轮廓；左下为窗口宽度 40，在灰度级范围 75 至 115 时的滤波结果，右下是相应的轮廓；

3.4 融合算法及轮廓选择

经过剪切算法后，我们会获得一系列候选轮廓 $C = \{C_p\}_{p \in S}$ 。这些轮廓可能已经包括了少数准确的建筑物轮廓，但是绝大多数轮廓都可能属于非建筑物轮廓或者只是建筑物本身一部分的轮廓。同时，在 C 中存在着大量相互重叠的轮廓区域。假设建筑物轮廓空间为 $B = \{b_i\}$ ，我们希望通过把 C 中相互重叠的轮廓区域相互融合获得新的轮廓候选空间： $C' = C \cup C^{merge}$ （ C^{merge} 表示融合后的轮廓构成的空间），使满足： $C' \supset B$ 。这样便可以删除掉那些非建筑物轮廓从而保留建筑物轮廓。在进行融合算法之前，可以进行轮廓预选择——利用一些简单的准则去掉部分非建筑物区域从而提高算法效率；在融合算法结束后使用特定的准则

将获得最终的建筑物轮廓。在这一结中，首先将描述轮廓预选择，然后将介绍融合算法和轮廓选择。

3.4.1 轮廓预选择

首先需要说明的是，此小结中介绍的准则同样可以作为轮廓的特征嵌入我们的概率模型中。之所以没有这样做，主要是基于算法效率的考虑：这些简单的特征很容易计算，直接使用它们可以快速删除那些非建筑物轮廓；同时可以减少模型中特征向量的维度，从而减少模型参数优化和轮廓概率值的计算开销。我们采用如下两个准则：

✓ 面积准则

对于那些区域面积很小的轮廓来说，他们对融合算法的贡献很少并且会增加可观的计算量，因此可以根据预先定义的最小轮廓区域阈值删除那些区域面积过小的轮廓；对于那些区域面积过大的轮廓来说，由于真实世界建筑物屋顶面积有限，所以可以根据预先定义的最大轮廓区域阈值删除那些区域面积过大的轮廓。

✓ 长度准则

计算每个轮廓的最小约束矩形，约束矩形的较长边长大于了预先给定长度阈值的轮廓可以被直接删除，这是基于现实世界中建筑物边界长度是有限的事实。

3.4.2 融合算法和轮廓选择

融合的目的是把相互重合的轮廓区域进行合并而获得新的轮廓区域。一般来说，即使是一幅很小的图像，经过剪切算法后也会获得很多候选轮廓，其中具有重叠区域的轮廓数量也很多，而且某个轮廓往往同时和其它多个轮廓相交。因此获得整个 C^{merge} 空间将是一个困难的组合问题，实际上很难实现。这里我们采用一个启发式融合方法进行轮廓合并以简化这个问题。我们假设已经建立了一个概率模型 $P_i^{building}$ 用以反映轮廓 C_i 是建筑物轮廓的概率——概率值越高则轮廓越有可能是建筑物轮廓。在 3.5 节中将具体给出这个概率模型。我们的融合方法定义如下：若轮廓 C_i 存在和它具有重叠区域的轮廓集合 $C^{inter} = \{C_j\}$ ，

$C^{inter} \subset C$ 。以 C_i 作为“种子”与轮廓 C_j 进行合并，合并后的轮廓为 C_{ij}^{merge} ，计算 C_{ij}^{merge} 的概率 $P_{ij}^{building}$ 。若此概率大于 C_i 的概率 $P_i^{building}$ ，则删除 C_i 和 C_j ，替换 C_i 为 C_{ij}^{merge} ；若此概率小于 C_i 的概率 $P_i^{building}$ ，则保留 C_i 并尝试与 C^{inter} 中下一个轮廓合并。重复以上算法直至遍历 C^{inter} 中所有轮廓，然后更新候选轮廓集 C ，删除 C 中所有和 C_i 相交的轮廓。反复使用以上介绍的融合算法直至整个候选轮廓集 C 中不再有重叠轮廓。最后在 C 中根据 $P_i^{building}$ 的大小使用特定的阈值 $THRESHOLD$ ，删除非建筑物轮廓，保留建筑物轮廓，得到最终结果。融合和选择算法的伪代码如下所示：

```

procedure MERGE_SELECT_PROCESS( $C$ )
    num: total contour number of set  $C$ 
    for  $i = 1$  to num
        for  $j = i + 1$  to num
            if  $C[i] \cap C[j] \neq \emptyset$ 
                 $C_{ij}^{merge} = C[i] \oplus C[j]$ 
                if  $P_{ij}^{building} > P_i^{building}$ 
                     $C[i] = C_{ij}^{merge}$ 
                endif
            endif
        repeat
            delete all the contours overlaid with  $C_{ij}^{merge}$  in  $C$ 
             $C = C \cup C_{ij}^{merge}$ 
        repeat
            for  $i = 1$  to num
                if  $P_i^{building} < THRESHOLD$ 
                    delete  $C[i]$ 
                endif
            repeat
        end MERGE_SELECT_PROCESS
    
```

算法中“ $\oplus$ ”符号的意义是把两个轮廓对应的区域进行合并，然后提取合并后区域的最外层闭合轮廓。

图 3-8 分别例示了经过剪切算法和融合算法后的结果。

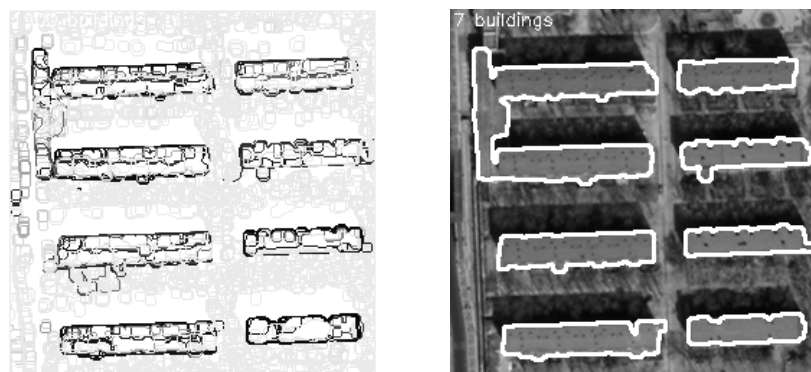


图 3-8 剪切算法和融合算法示例。左图是经过剪切算法后得到的轮廓，其中灰度值越低则轮廓建筑物类概率越高；右图是经过融合算法和轮廓选择后得到的最终建筑物轮廓在原始图像上的叠加。

尽管我们使用的启发式融合算法在大多数情况下都能够获得较好的结果。但是由于这种算法仍然无法遍历相交轮廓所有可能的融合情况，因此具有一定的局限性，主要表现在两方面：

- ✓ 融合顺序的选择：目前的算法没有考虑融合顺序对融合结果的影响。比如轮廓 C_i 和 C_j 相交，融合后的轮廓为 C_{ij}^{merge} ，且 $P_{ij}^{building} > P_i^{building}$ ；同时 C_i 也和 C_p 相交，融合后的轮廓为 C_{ip}^{merge} ，且有 $P_{ip}^{building} > P_i^{building}$ 。在算法中，我们可能先把 C_i 和 C_j 进行融合，也可能先把 C_i 和 C_p 进行融合，这两种情况将导致分别以不同的新的种子点 C_{ij}^{merge} 和 C_{ip}^{merge} 继续进行算法，显然，两种情况下最终算法融合的结果也会不同。这种由于融合顺序不同带来的不确定性可能会导致融合的最终结果并不是最好的融合结果，而可能是次好的结果。但是，就目前来说，没有一种好的方法能够高效的遍历所有的融合情况，选用启发式搜索忽略融合顺序是一种折衷的选择。
- ✓ 轮廓间拓扑结构的影响：我们的算法中没有考虑轮廓间拓扑结构的影响。比如，我们可能直接由剪切算法或者在融合算法进行过程中，得到建筑物的子结构区域轮廓——比如屋顶的阁楼，这些子结构相对于建筑物区域来说面积很小，但是由于具有很强的建筑物特征，会获得很高的 $P^{building}$ 值，这样在继续进行融合算法时这个子结构轮廓就可能被保留下来，而无法得到整个建筑物的轮廓。

3.5 模型建立

由于我们需要寻找的是建筑物轮廓，因此轮廓候选集分为两类：建筑物轮廓与非建筑物轮廓。我们希望使用一个二类概率分布模型来衡量每一个轮廓属于建筑物类或者非建筑物类的可能性。在这里，我们借鉴了 Kumar 提出的 DRF 模型[46]中的联合势，定义我们自己的模型为：

$$P_i = P(x_i | y_i) = \sigma(x_i \omega^T f(y_i)) = \frac{1}{1 + e^{x_i (\omega^T f(y_i))}} \quad \text{公式 3-1}$$

其中 i 表示轮廓候选集 C 中第 i 个轮廓 C_i ； y_i 是从轮廓 C_i 中提取的一系列特征构成的向量。 x_i 是轮廓属于某一类别的标记：当轮廓属于建筑物轮廓时 $x_i=1$ ，属于非建筑物轮廓时 $x_i=-1$ 。 ω 是模型参数向量。一般来说，通过提高特征空间的维度能获得更准确的分类平面[47]，因此我们希望通过提升特征空间 y 的维度，提高整个模型的分类能力；同时，越高阶的特征空间对于噪声也就越敏感。综合考虑，我们利用特征空间转换函数 $f(\bullet)$ 把 y 从线性特征空间 $\mathfrak{R}^m$ 转换到二次特征空间 $\mathfrak{R}^d$ （ m 和 d 为特征空间的维度）[47]；即 $y \rightarrow [1, y, y^t y, \tilde{y}]$ ，其中 $\tilde{y}$ 是 y 中特征值两两相乘构成的向量。若已知 y 的维度为 m ，则相应的二次特征空间的维度 $d = 1 + m * (m + 3) / 2$ 。 $\sigma(\bullet)$ 是逻辑函数，其函数曲线如图 3-4 所示。

根据逻辑函数的性质，很容易证明：

$$\sum_x P(x) = P(x=1) + P(x=-1) = 1 \quad \text{公式 3-2}$$

若已知参数 ω 及轮廓 C_i 的特征向量 y_i ，根据公式 3-1 可以计算出 C_i 建筑物类与非建筑物类的概率值。若轮廓满足 $P(x=1) > P(x=-1)$ ，则认为 C_i 是建筑物轮廓；反之则认为 C_i 是非建筑物轮廓。或者简单地，由公式 3-2，我们根据 $P(x=1) > 0.5$ 可以直接判断轮廓是否为建筑物轮廓。

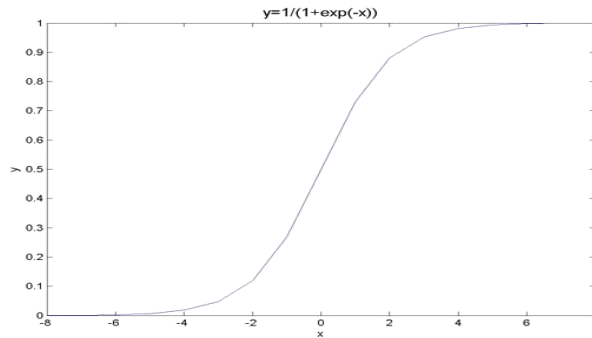


图 3-4 逻辑函数曲线

3.6 模型参数优化

若模型参数 ω 是未知的，则模型便不具备分类能力，需要利用先验知识对模型参数进行优化。我们预先给定一系列图像及相应图像中的建筑物轮廓和非建筑物轮廓以构成训练集。假设训练集中图像个数为 N ， S 表示每幅图像中建筑物轮廓 ($x_i=1$) 和非建筑物轮廓 ($x_i=-1$) 构成的轮廓空间。假设图像 n 给定 $\mathbf{y}^n$ 时 x_i^n 之间相互独立，整个训练集中所有轮廓的概率分布可表示为：

$$P = \prod_n \prod_{i \in S} P(x_i^n | \mathbf{y}_i^n) \quad \text{公式 3-3}$$

我们采用“最大似然逻辑回归”(maximum-likelihood logistic regression) [48] 对参数 ω 进行估计：

$$\omega_{opt} = \arg \max_{\omega} \sum_n \sum_{i \in S} \log P(x_i^n | \mathbf{y}_i^n) \quad \text{公式 3-4}$$

其中对数似然函数 (log-likelihood) 为：

$$L(\omega) = \sum_n \sum_{i \in S} \log P(x_i^n | \mathbf{y}_i^n) \quad \text{公式 3-5}$$

其二阶导数矩阵 (Hessian 矩阵 ($d \times d$ 维)) 为：

$$\nabla_{\omega}^2 L(\omega) = - \sum_n \sum_{i \in S} \left\{ \sigma(\omega^T \mathbf{f}(\mathbf{y}_i)) (1 - \sigma(\omega^T \mathbf{f}(\mathbf{y}_i))) \right\} \mathbf{f}(\mathbf{y}_i) \mathbf{f}^T(\mathbf{y}_i) \quad \text{公式 3-6}$$

根据公式 3-6，Hessian 矩阵与轮廓的标记 x_i 无关且非正，因此公式 3-5 是凸的，任何局部最大值都是全局最大值。针对以上特性，我们可以采用较为高效的 Newton 梯度下降法 [48] 最大化公式 3-4，得到最优参数 ω_{opt} 。

3.7 特征的选择

本结介绍如何计算轮廓的特征，即向量 $\mathbf{y}$ 的构成。首先介绍在计算轮廓特征时需要预先计算的特征图像，然后将详细描述各个轮廓特征的计算方法。

3.7.1 特征图像

在我们的算法中，需要预先计算四种二元特征图像：边缘图像，直线段图像，直线段距离图像和阴影遮罩。它们分别应用于不同的轮廓特征计算中。下

面分别描述它们的计算方法。图 3-5 例示了这些特征图像。

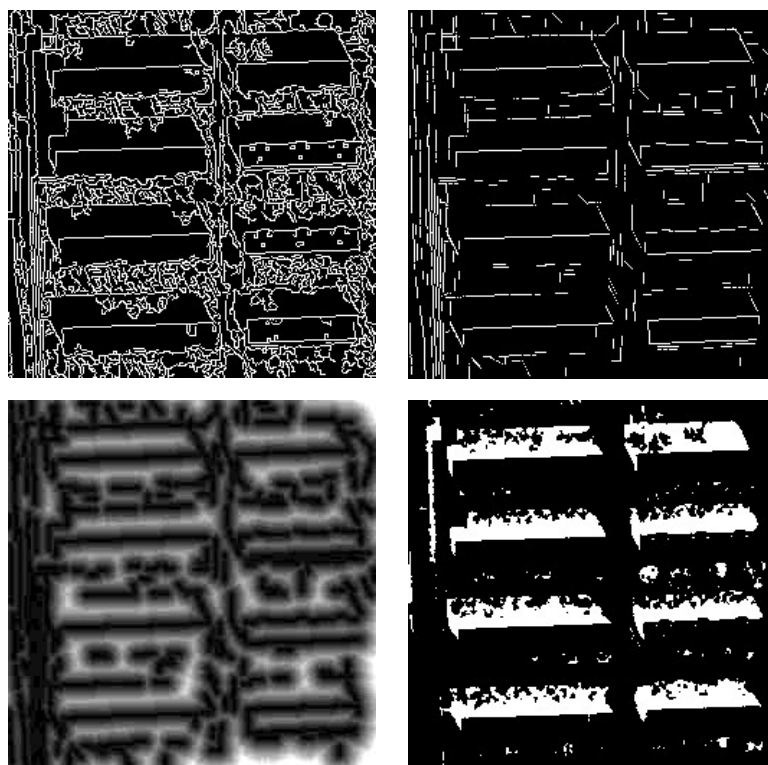


图 3-5 图 3-2 图像的特征图像。左上为边缘图像；右上为直线段图像；左下为直线段距离图像；右下为阴影遮罩。

✓ 边缘图像

边缘图像主要描述输入图像中的边缘信息，即图像中像素灰度变化剧烈的像素点。我们采用 Canny 边缘检测算子进行计算。

✓ 直线段图像

直线段图像描述输入图像中构成直线的像素点，即梯度方向一致的像素点。我们采用[49]中的方法基于边缘图像找出其中的直线段。

✓ 直线段距离图像

我们根据直线段图像计算直线段距离图像，具体算法为计算图像中各个像素点到最近的直线段像素点的距离，并把这个距离保存为此像素点的像素值。

✓ 阴影遮罩

阴影遮罩记录图像中阴影区域的像素点，即那些灰度级较低的像素点。具体算法为首先计算输入图像灰度直方图，然后进行聚类并找到直方图中灰度级最低的那一类作为图像中的阴影类。阴影类的灰度阈值表示为 SHADOW_THRESHOLD。

3.7.2 轮廓特征

我们关注城市区域中建筑物的提取，因此我们需要针对这个特点寻找有效的建筑物特征。在城市区域中，存在着大量密集的人造物体和自然物体：比如建筑物、道路、公园、水体，植被等等；它们之间各自具有不同的特征，同时也具有相似的特征。对于大部分人造物体来说，它们具有很多共同的特征：比如物体边缘主要由直线构成，形状为规则几何多边形，纹理规则均匀等等。通过这些特征可以把绝大部分人造物体和自然物体区分开，但为了进一步识别出建筑物，还需要更多更准确的仅属于建筑物的特征，比如高度特性等。由于本算法是基于轮廓这个基本元素的，因此本算法中的特征也是基于轮廓或轮廓区域的特征。下面将描述算法中使用的八个特征：其中一部分是人造物体共同的特征，另一部分是建筑物独有的特征。

✓ 轮廓像素与直线的匹配度

大部分建筑物轮廓都可以近似为多边形，即由直线段构成。因此，我们自然想到计算轮廓和直线段的关系。可以简单的把某个轮廓和预先计算的直线段图像进行叠合，计算匹配点的数目，以此反映轮廓和直线段的匹配程度。但是这里有一个问题：当进行参数训练时，建筑物轮廓可能来自于人工手绘或者 GIS，不管采用哪种方式，都很难保证作为先验知识的建筑物轮廓和图像中真实建筑物轮廓像素点准确匹配，这就会影响特征计算的准确度，进而影响模型参数优化的结果。鉴于上述原因，我们采用鲁棒的方法：即计算轮廓和图像中直线段像素点的平均距离。假设轮廓 C_i 由 EDGE_NUM 个像素点构成，我们利用预先计算的直线段距离图像计算轮廓 C_i 对应点的像素值的和：SUM_LINE_DIS——即 C_i 每个像素点到图像中最近直线段距离的和。则轮廓 C_i 到图像中直线段像素点的平均距离：

$$AVG_LINE_DIS = SUM_LINE_DIS / EDGE_NUM \quad \text{公式 3-7}$$

AVG_LINE_DIS 的值越小则与轮廓接近或重合的直线段像素点越多，轮廓越接近

多边形。

✓ 轮廓区域内包含的非直线边缘

一般来说，建筑物的屋顶可能是平滑的或者含有少量的直线段；而对于非建筑物轮廓，其轮廓区域内一般存在大量非直线边缘。因此我们计算轮廓区域内非直线边缘的比率——INCLUDE_NON_LINE_EDGE 来反映建筑物轮廓的这一特征。假设轮廓 C_i 的内部区域 Λ_i 内含有的像素点数目为 IN_NUM，利用预先计算的边缘图像计算 Λ_i 内含有的边缘像素点数目 IN_EDGE_NUM，利用预先计算的直线段图像计算 Λ_i 内含有的直线段像素点数目 IN_LINE_NUM，则 INCLUDE_NON_LINE_EDGE 可表示为：

$$\text{INCLUDE_NON_LINE_EDGE} = (\text{IN_EDGE_NUM} - \text{IN_LINE_NUM}) / \text{IN_NUM} \quad \text{公式 3-8}$$

INCLUDE_NON_LINE_EDGE 的值越小则轮廓 C_i 内包含的非直线段边缘越少，轮廓越有可能是建筑物轮廓。

✓ 轮廓区域灰度熵值

大部分建筑物屋顶都是平滑的或者由简单的纹理构成，而非建筑物轮廓则正好相反。在信息论中，“熵”可以衡量信息量的多少[50]。基于区域的像素灰度的熵反映这一区域内像素灰度信息量的多少。我们基于区域 Λ_i 内像素灰度直方图计算这一区域内的像素灰度的熵——IN_ENTROPY。IN_ENTROPY 的值越大表明轮廓区域内的像素信息量越小，轮廓越有可能是建筑物轮廓。

✓ 轮廓区域灰度中值和标准差

一般来说，绝大部分建筑物屋顶区域的灰度平均值要高于阴影区域的像素灰度。因此，对于轮廓区域 Λ_i 来说，其像素灰度平均值能够区分开阴影区域轮廓和非阴影区域轮廓。在我们的算法中，我们采用像素灰度中值——IN_MEAN取代平均值，能够获得更好的鲁棒性[51]。我们定义IN_MEAN₂：

$$\text{IN_MEAN}_2 = \begin{cases} 0 & \text{if } \text{IN_MEAN} > \text{SHADOW_THRESHOLD} \\ \text{IN_MEAN} / (\text{SHADOW_THRESHOLD}) & \text{if } \text{IN_MEAN} < \text{SHADOW_THRESHOLD} \end{cases} \quad \text{公式 3-9}$$

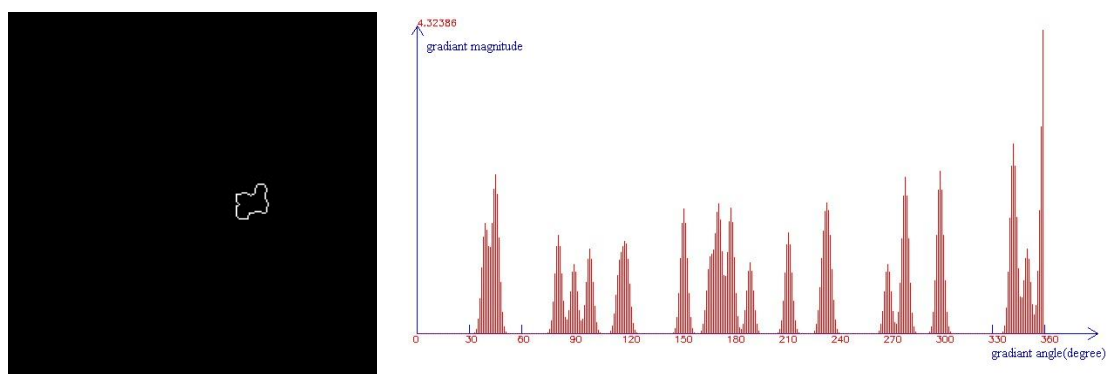
由于建筑物本身材质及其反射率的多样性，因此在图像上它们可能取任何的灰

度值，所以在 $IN_MEAN > SHADOW_THRESHOLD$ 时，设置常数 0 消除 IN_MEAN_2 这一特征的作用；另一方面，由于阈值 $SHADOW_THRESHOLD$ 并不能保证把阴影区域和非阴影区域完全分开，且灰度值越低的区域越有可能是阴影区域，所以 $IN_MEAN < SHADOW_THRESHOLD$ 时，我们采用比值作为特征值。

一般来说，绝大部分建筑物屋顶区域的灰度分布标准差 IN_DEV 不会太大， IN_DEV 在一定程度上可以区分开一些建筑物与非建筑物轮廓区域。因此，我们把 IN_DEV 作为特征向量中的一个特征。

✓ 轮廓的形状特征

虽然建筑物的形状多种多样，但是在我们的算法中，我们认为建筑物的轮廓都是由相互垂直或者平行的直线构成的，这种假设是符合大多数实际情况的。为了描述这个特征，我们计算轮廓像素点的梯度方向和梯度大小，以梯度方向为横坐标，梯度大小为纵坐标计算轮廓的梯度直方图。然后计算直方图的二阶矩—— $SHAPE_MOMENT$ 并把其作为一个特征，这个特征反映了直方图中各个梯度方向的梯度大小高于梯度平均值的程度。若 $SHAPE_MOMENT$ 较大，则表明轮廓像素点的梯度方向集中于少数几个确定的方向上，即轮廓可能是由几条具有确定方向的直线构成；反之轮廓则可能由一些曲线构成。因此矩特征可以区分开多边形轮廓与曲线轮廓。另外，假设 δ_1 和 δ_2 分别表示梯度直方图中的第一大和第二大梯度峰值对应的梯度方向，因此基于方向的形状特征—— $SHAPE_ANGLE$ 可表示为 $|\cos(2(\delta_1 - \delta_2))|$ 。这个特征反映了轮廓主要两个边的平行性或者垂直性： $SHAPE_ANGLE$ 越大则这种性质越突出。图 3-6 例示了这两个基于形状的特征。



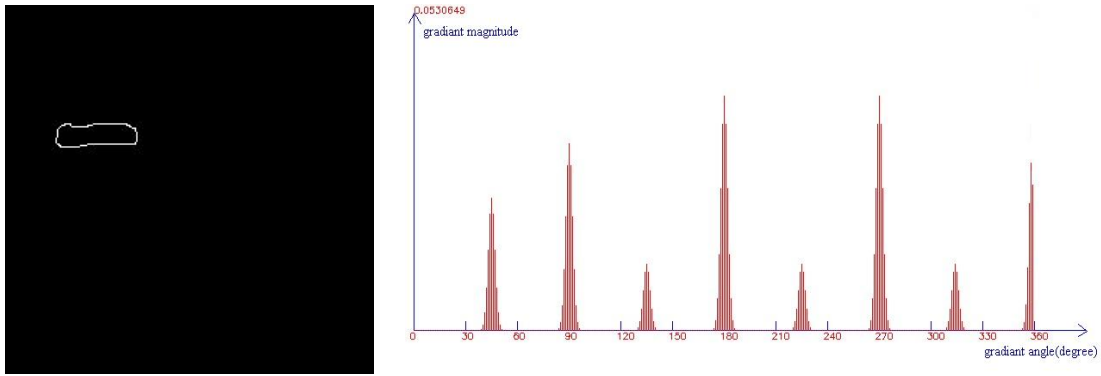


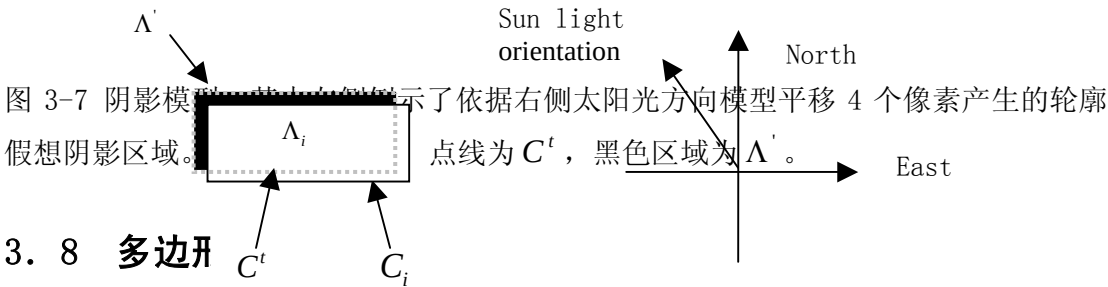
图 3-6 形状特征。左上是一个不规则形状，右上对应它的直方图，SHAPE_MOMENT = 5.23, SHAPE_ANGLE = 0.56。左下是相对规则形状，右下对应它的直方图，SHAPE_MOMENT = 20.2, SHAPE_ANGLE = 1.0。

✓ 轮廓与阴影的关系

高度是建筑物区别于其它人造物体的显著特征。在一般的高分辨率遥感光学图像中，建筑物的高度是通过其墙面和阴影反映出来的。由于 QUICKBIRD 图像本身的分辨率有限，并且拍摄角度近乎垂直，很难利用建筑物的墙面信息，因此，我们求助于阴影信息。我们认为只有和阴影区域紧密相联的轮廓区域才可能是建筑物的轮廓区域，因此定义一个名为“轮廓阴影率”（SHADOW_RATE）的特征来反映这个特性。我们很容易通过简单的人工计算获得图像成像时的阳光照射角度，可以推算出物体可能投射阴影的区域：依据阳光照射角度，我们移动 C_i 一小段距离（3-5 像素），假设移动后的轮廓为 C^t 。它们的内部区域分别表示为 Λ_i 和 Λ^t ，则轮廓假想阴影区域 Λ' 可表示为 $\Lambda' = (\Lambda_i \cup \Lambda^t) \setminus \Lambda_i$ ，如图 3-7 所示。根据预先计算的阴影遮罩，计算区域 Λ' 中包含的阴影像素点的数目——IN_SHADOW_NUM 并假设区域 Λ' 中像素点数目为 IN_NUM，则特征 SHADOW_RATE 可表示为：

$$\text{SHADOW_RATE} = \text{IN_SHADOW_NUM} / \text{IN_NUM} \quad \text{公式 3-10}$$

SHADOW_RATE 表示轮廓假想阴影区域中存在的实际阴影像素点的多少，其值越高则轮廓越有可能是建筑物轮廓。



3.8 多边形

获得了建筑物的轮廓后，可以看到这些轮廓形状并不规则，轮廓只能大体上反映真实建筑物的形状，仍存在诸如小凸起或者小凹陷之类的噪声；另外，一般建筑物识别算法的结果都是为以后的操作或其他算法进行服务的（比如在地图更新中用以更新 GIS 数据），它们往往需要矢量化数据格式。综上，我们希望用一种矢量化的描述代替像素集合的描述方式，并平滑轮廓形状，从而把建筑物轮廓转化为更易于描述的矢量多边形。这就是多边形近似的目的和意义。

我们使用 Douglas - Peucker 算法[52]对提取出的建筑物轮廓进行多边形近似。可以控制近似多边形和轮廓的近似程度从而找到较好的多边形近似结果。图 3-9 例示了图 3-8 中轮廓的近似结果。

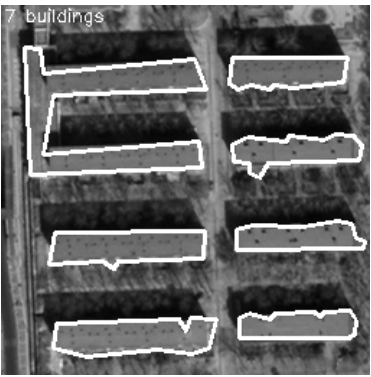


图 3-9 多边形近似结果。近似度为 0.1。

3.9 实验结果及分析

这一节将对我们提出的算法进行验证及分析，首先介绍实验中算法参数设置，然后将展示试验结果并作以分析。我们选用的数据是 2002 年和 2003 年北京地区 QUICKBIRD 全色卫星图像，分辨率为 0.6 米/像素。图像数据格式为单通道 8 比特深度位图(BMP)图像。

模型参数优化需要提供训练集，其中包括图像数据和其对应的建筑物轮廓数据和非建筑物轮廓数据。图像数据选取 5 幅 2002 年的图像，大小在 1200X1200 左右。图像中对应的建筑物轮廓可以采用两种方式获得：一种是通过人工手绘，这种方法不存在配准问题，并且可以得到比较准确的建筑物轮廓从而获得较好的优化参数，缺点是费时费力；另一种是采用 GIS 作为轮廓数据，其优点是效率高，缺点是由于 GIS 和图像存在配准问题及 GIS 本身准确度的局限会导致参数优化的质量下降。目前我们只具有 2001 年的 GIS 数据，与 2002 年的 Quickbird 无法匹配，因此试验中使用的训练集由第一种方法生成。非建筑物轮廓数据可以通过在图像中使用剪切算法寻找建筑物轮廓区域之外区域中的轮廓获得。在实验中，我们使用了 700 个建筑物轮廓和 2100 个非建筑物轮廓，之所以采用更多的非建筑物轮廓数据是因为非建筑物的形态更多样。模型参数优化的计算时间大概为 8 分钟，其中大部分时间用于计算轮廓特征，优化算法计算时间少于 10 秒。

3.9.1 参数设置

实验中算法各个步骤的参数设置如下所示：

✓ 剪切算法参数

我们采用 5 级多滑动窗口二元带通滤波器。由于融合算法会将重叠区域进行融合，所以不需要设置过大的窗口长度，同时考虑到效率问题，我们设置各个窗口的大小分别为：10，20，30，40，50，步长为 1。滤波起始灰度级我们选择生成阴影遮罩图像时得到的阴影阈值 SHADOW_THRESHOLD，灰度级上限我们选择 8 比特图像的上限 255。在进行数学形态学算子去噪声时，我们使用的是十字型的算子元素。

✓ 特征参数

在计算轮廓特征时，我们需要预定义的参数有两个：一个是图像成像时的太阳光近似方位角；一个是计算轮廓阴影率时轮廓平移的距离。我们所用 2002 年数据的太阳光方位角大致为北偏西 30 度。轮廓平移距离设置为 4 像素。在计算轮廓特征所需的 Canny 图像时，选择 Canny 滤波器上下阈值为 50 和 200。

✓ 融合与选择参数

轮廓预选则时，定义轮廓面积的下限为 200 像素，上限为 15000 像素。约束矩形较长边的边长上限为 400 像素。

✓ 多边形近似参数

Douglas - Peucker 算法中的近似参数，即近似多边形和轮廓的周长比定义为 0.85。

3. 9. 2 结果及分析

试验使用 9 幅图像作为测试集，它们同样来自于 2002 年，并保证与训练集图像没有交集。我们参考 Wenbo Song 今年提出的关于建筑物提取结果评价准则的描述[2]，并结合我们算法的具体情况，采用如下 7 个评价标准。其中“ref”下角标表示用于参考的真实数据（ground truth），是由人工绘制的。

■ 基于数量的评价：

假设真实建筑物数量为 NUM_{ref} ，算法识别出的建筑物轮廓数量为 NUM_{det} ，其中正确的建筑物轮廓数量为 NUM_{det_right} ，错误的数量为 NUM_{det_error} ，未检测到的数量为 NUM_{det_non} ，则：

✓ 识别率（NumDetR）： $NumDetR = NUM_{det_right} / NUM_{ref}$

✓ 正确率（NumTrueR）： $NumTrueR = NUM_{det_right} / NUM_{det}$

■ 基于区域面积的评价：

假设所有真实建筑物轮廓区域的面积和为 $AREA_{ref}$ ，算法识别出的所有建筑物轮廓区域的面积和为 $AREA_{det}$ ，算法识别出的建筑物轮廓区域与真实建筑物轮廓区域的公共区域的面积和为 $AREA_{inter}$ ，则：

- ✓ 面积正确率 (AreaTureR): $\text{AreaTureR} = \text{AREA}_{\text{inter}} / \text{AREA}_{\text{ref}}$
- ✓ 面积错误率 (AreaErrorR): $\text{AreaErrorR} = (\text{AREA}_{\text{det}} - \text{AREA}_{\text{inter}}) / \text{AREA}_{\text{det}}$
- ✓ 面积遗失率 (AreaLoseR): $\text{AreaLoseR} = (\text{AREA}_{\text{ref}} - \text{AREA}_{\text{inter}}) / \text{AREA}_{\text{ref}}$

■ 基于轮廓形状的评价:

假设所有真实建筑物轮廓区域的周长和为 PERI_{ref} , 算法识别出的所有建筑物轮廓区域的周长和为 PERI_{det} , 则:

- ✓ 周长相似度 (PeriR): $\text{PeriR} = \text{ABS}(\text{PERI}_{\text{ref}} - \text{PERI}_{\text{det}}) / \text{PERI}_{\text{ref}}$
- ✓ 面积相似度 (AreaR): $\text{AreaR} = \text{ABS}(\text{AREA}_{\text{det}} - \text{AREA}_{\text{ref}}) / \text{AREA}_{\text{ref}}$

表 3-1 和表 3-2 展示了这些评价标准应用于测试集的结果。图 3-10 至图 3-12 例示了训练集中三幅图像建筑物轮廓提取结果。从三幅图中可以看到我们试验图像场景具有较大复杂度, 具体表现在场景中物体比较密集且建筑物轮廓本身纹理和形状具有一定的多样性。在这样的数据集中, 我们的算法均能取得较好的结果, 其中概率模型和剪切-融合算法发挥了巨大的作用。概率模型相当于算法的“大脑”, 剪切-融合算法相当于算法的“双手”——概率模型综合了训练集给出的先验知识, 能够告诉“双手”什么样的轮廓才是建筑物轮廓; 而“双手”则负责拼凑挑选出那些符合要求的轮廓。从表 3-1 中可以看出, 此算法具有较高的识别率和正确率, 并且正确率高于识别率, 这是由于算法提取出的建筑物轮廓中“纯粹”的非建筑物轮廓非常少, 这也反映了算法具有很强的剔除非建筑物轮廓的能力。在计算识别率与正确率时, 只考虑了轮廓的数量——即只要轮廓和真实建筑物区域重合便认为此轮廓为检测到的正确建筑物轮廓, 所以我们需要用区域面积来衡量轮廓区域的准确性。从表 3-2 中可以看到面积的错误率较低而遗失率较高, 说明了算法寻找到的冗余的错误区域较少但丢失的建筑物区域较多。这是由于在进行剪切算法时, 我们会运用数学形态学去噪声, 此时不可避免地会丢失一些较小的非噪声区域; 同时在融合算法中, 可能由于模型参数或者特征的原因, 某些本属于建筑物部分区域的区域在和“种子”区域合并后并不能产生更大的概率值, 它们也就没有被融合而保留下来。两个形状特征的值都相对偏高, 表明建筑物轮廓的形状准确度并不好, 这可能是来自于两方面的原因: 一是我们的算法并不像传统的建筑物提取算法那样是基于角点或者基于图像上直线段进行聚类的, 而是采用轮廓这一基本元素。而轮廓边

缘一般都具有很多噪声，形状并不规则，而真实建筑物轮廓往往是规则的多边形，这就造成了形状上的差异；另一个原因可能来自于多边形近似，由于近似算法本身的局限，不可能得到完美的接近于真实建筑物形状的多边形，并且近似的过程也带来一定的误差，从而影响最终的形状相似性。

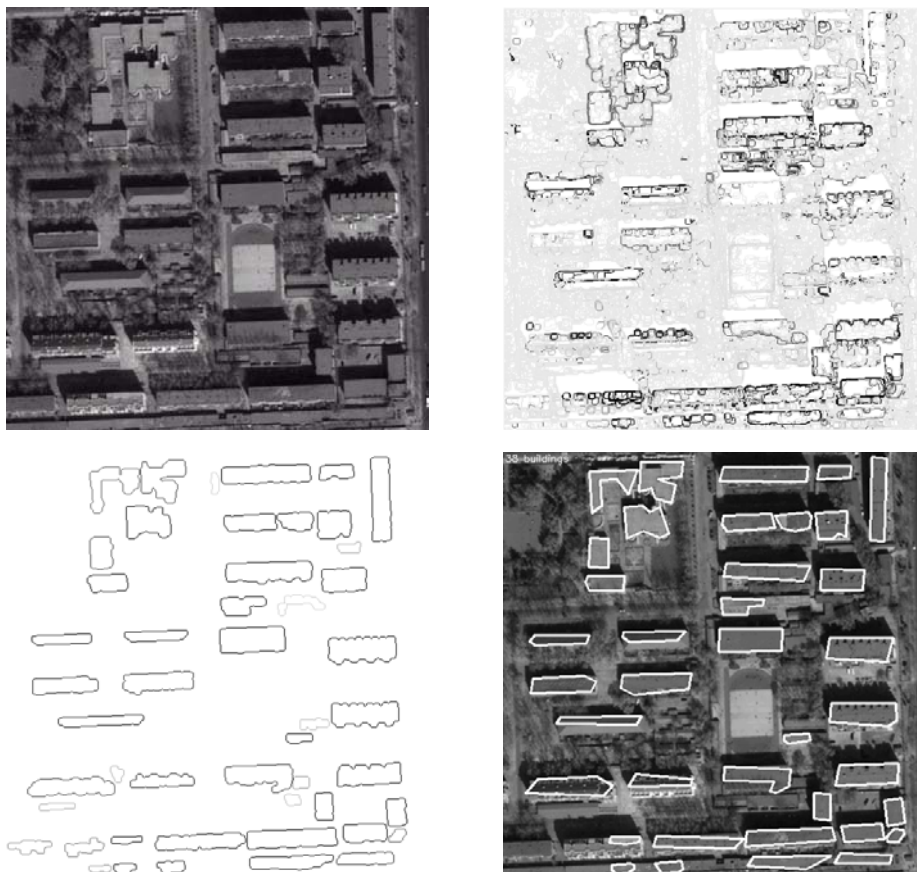


图 3-10 1 号图像建筑物轮廓识别。左上为原始图像；右上为剪切算法结果；左下为融合算法结果，灰度值越低则轮廓建筑物类概率越高；右下为多边形近似结果在原始图像上的叠加。图像大小：512x512。算法计算时间：4 分钟。

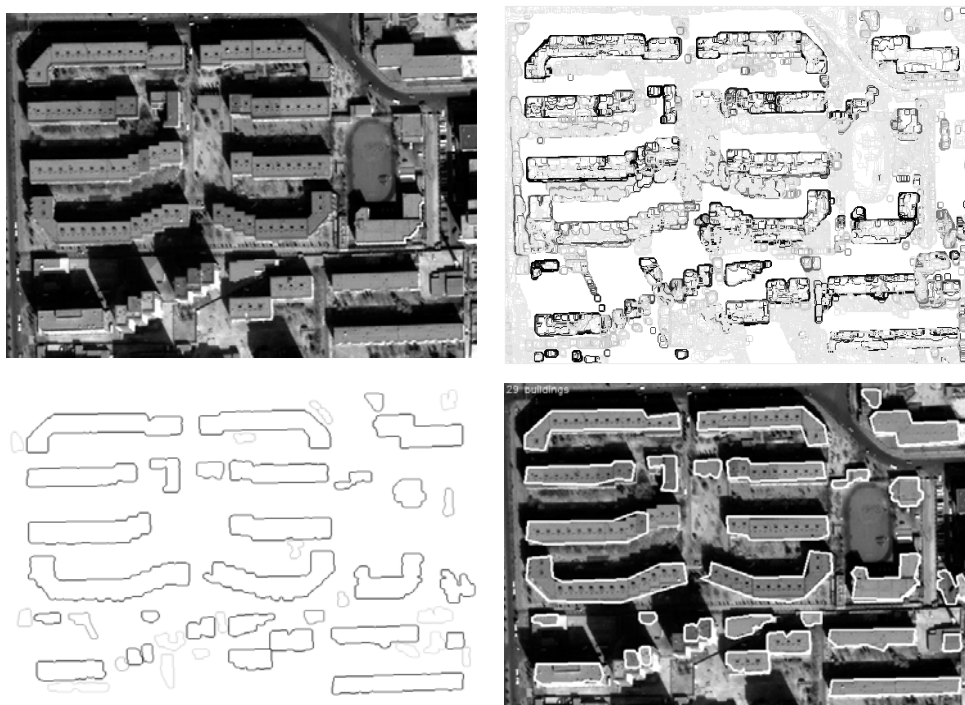
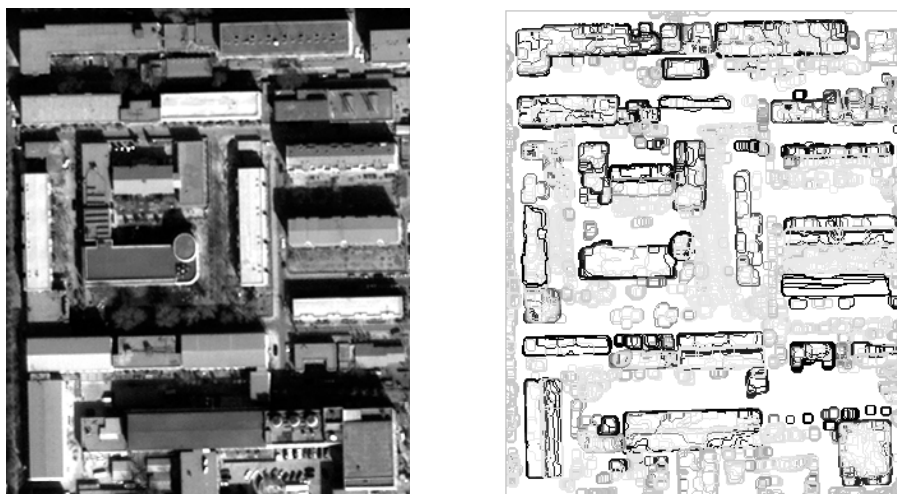


图 3-11 4 号图像建筑物轮廓识别。图像意义见图 3-10 注释。图像大小：574x408。算法计算时间：4 分钟。



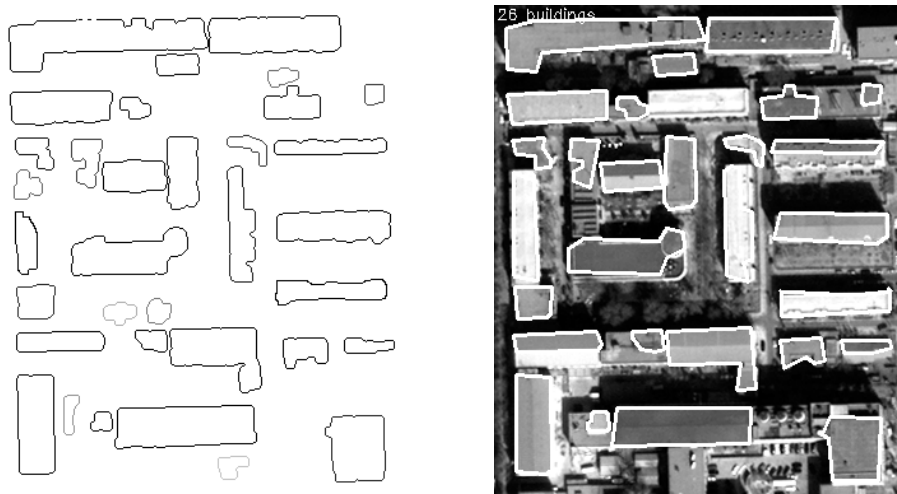


图 3-12 9 号图像建筑物轮廓识别。图像意义见图 3-10 注释。图像大小：316x384。算法计算时间：3 分钟。

Image	NUM _{ref}	NUM _{det_right}	NUM _{det_error}	NUM _{det_non}	NumDetR	NumTrueR
1	37	36	0	1	0.973	1.000
2	30	24	1	6	0.800	0.960
3	40	34	3	6	0.850	0.919
4	24	24	2	0	1.000	0.923
5	35	24	2	11	0.686	0.923
6	23	20	2	3	0.870	0.909
7	16	12	1	4	0.750	0.923
8	44	34	4	10	0.773	0.895
9	28	23	1	5	0.821	0.958
Total	277	231	16	46	0.834	0.935

表 3-1 基于轮廓数量的评价

Image	AreaTureR	AreaErrorR	AreaLoseR	PeriR	AreaR
1	0.902	0.035	0.098	0.226	0.133
2	0.890	0.070	0.110	0.198	0.199
3	0.811	0.053	0.189	0.301	0.251
4	0.923	0.049	0.077	0.241	0.192
5	0.815	0.068	0.185	0.290	0.184
6	0.887	0.072	0.113	0.155	0.121
7	0.832	0.059	0.168	0.251	0.180
8	0.844	0.067	0.156	0.202	0.156
9	0.823	0.038	0.177	0.260	0.221
Total	0.861	0.060	0.139	0.244	0.189

表 3-2 基于轮廓面积和轮廓形状的评价

3. 10 小结

在这一章中，我们介绍了一种运用概率模型在单幅高分辨率卫星图像上提取建筑物的算法。这种算法的核心和主要思想是概率模型和剪切-融合算法。概率模型提供了一种广泛的适应性——即能够同时结合多种特征并通过参数学习找到这些特征之间的联系，是一种从高层到低层的思想。同时概率模型具有很好的扩展性：如果能获得新的特征，比如光谱特性，高度信息（DEM 图像）等，能够很方便的加入到模型中来，只要再进行一次参数训练便能够获得更强有力的模型，从而得到更好的建筑物轮廓识别结果。剪切-融合算法蕴含底层到高层的思想——由于建筑物轮廓区域的灰度分布并不总是具有平滑性或连续性，简单的阈值很难获取建筑物区域；同时平滑性和连续性是多数基本算法最容易处理的情况。这样我们就通过剪切算法首先获得相对容易处理的平滑连续区域，利用融合算法获得更复杂更高阶的处理对象。概率模型和剪切-融合算法的联合使用结合了高层到低层算法和底层到高层算法的优点，能够产生更好的结果。

同时也应该看到，我们提出的算法仍然存在一些问题。具体表现在：对于现代化的较复杂的建筑物，算法的识别效果会下降；在理想情况下，应该采用来自于不同时间或空间的数据分别作为训练集和测试集，但是由于我们的数据

十分有限，目前我们只能在一幅大图像上分割训练集和测试集。考虑到特征中使用了很多有关灰度的特征，对图像的光照条件具有一定的要求，所以若测试集与训练集来源于不同时期的图像，识别效果也会有一定的下降。其次，目前选用的多边形算法并不十分理想；鉴于以上问题，我们可以看到以后工作的一些方向：尝试更具有分辨率的特征；采用其它方法寻找更合理的特征空间，比如PCA等；寻找更好的融合算法，能够遍历更多 C^{merge} 中的元素；利用更大的训练集并提高使用GIS进行参数训练的效果；寻找更好的多边形近似方法等等。

第四章 基于轮廓提取的建筑物变化检测算法

4.1 引言

上一章介绍了利用概率模型并进行建筑物轮廓识别的算法，我们希望把它应用于城市区域多时相图像建筑物变化检测上来。这一章我们将介绍一种基于概率模型的建筑物变化检测算法。这种算法的主要思想是利用第三章中的算法首先找出不同时相图像中的建筑物轮廓，然后利用我们提出的概率模型比较对应的建筑物轮廓并给出它们变化的可能性，最后根据建筑物相应的变化概率值判断它们是否发生了变化。

一般，我们说建筑物发生了变化，往往可能包括了三种情况：一是旧建筑物消失——即在旧图像中的建筑物在新图像中对应位置消失了；二是新建筑物出现——即在新图像中出现的建筑物在旧图像上对应位置没有出现；三是建筑物形状发生了变化——即旧图像和新图像同样位置的建筑物的形状不相同。根据第三章的算法特点和分析可以知道，我们还不能找出完全准确的建筑物轮廓，因此很难判断不同时相图像中相应建筑物的形状是否真正发生了变化。所以，这一章所介绍的算法仅仅处理前两种建筑物变化的情况——即找到那些新出现的建筑物和消失的旧建筑物。

一般情况下，两幅不同时相图像的成像条件不同，而且同一地区的环境可能发生变化，即使是同一建筑物本身在灰度级、纹理等方面也可能发生变化；另外我们的建筑物轮廓识别算法往往无法找到完全准确的建筑物轮廓。这样就可能造成某些建筑物轮廓提取不完整，或者具有非建筑物区域的冗余，也可能一个建筑物轮廓被拆分成几个小区域轮廓等情况。以上这些因素大大增加建筑物变化检测的难度。

下面，将首先介绍算法流程，然后将分别对流程中各个步骤作以详细描述，最后，将给出合成图像和真实图像的试验结果并进行分析。

4.2 算法流程

假设我们两幅不同时间拍摄的同一地区的图像 I^1 和 I^2 ，并且它们的获取时间满足关系： $t_1 < t_2$ 。由于图像成像条件不同，我们首先需要把两幅图像进行配准，配准后两幅图像中公共区域为 A_{match} 。假设 I^1 和 I^2 中区域 A_{match} 内的建筑物轮廓集合可分别表示为 $B_1 = \{b_i^1, i \in S_1\}$ 和 $B_2 = \{b_j^2, j \in S_2\}$ ，其中 b 表示建筑物轮廓；它们所对应的区域分别表示为 $A_1 = \{\Lambda_i^1, i \in S_1\}$ 和 $A_2 = \{\Lambda_j^2, j \in S_2\}$ ，其中 Λ 表示区域。 B_1 和 B_2 中没有变化的建筑物集合和变化的建筑物集合分别表示为： B_{non_change} 和 B_{change} 。只考虑两种建筑物的变化——消失和新出现，对应的集合为： $B_{destr.}$ 和 B_{new} 。于是我们有关系： $B_{change} = B_{destr.} \cup B_{new}$ ， $B_1 = B_{non_change} \cup B_{destr.}$ ， $B_2 = B_{non_change} \cup B_{new}$ 。进行建筑物的变化检测即找出集合 B_{non_change} ， $B_{destr.}$ 和 B_{new} 。在理想情况下，我们有如下判决准则：

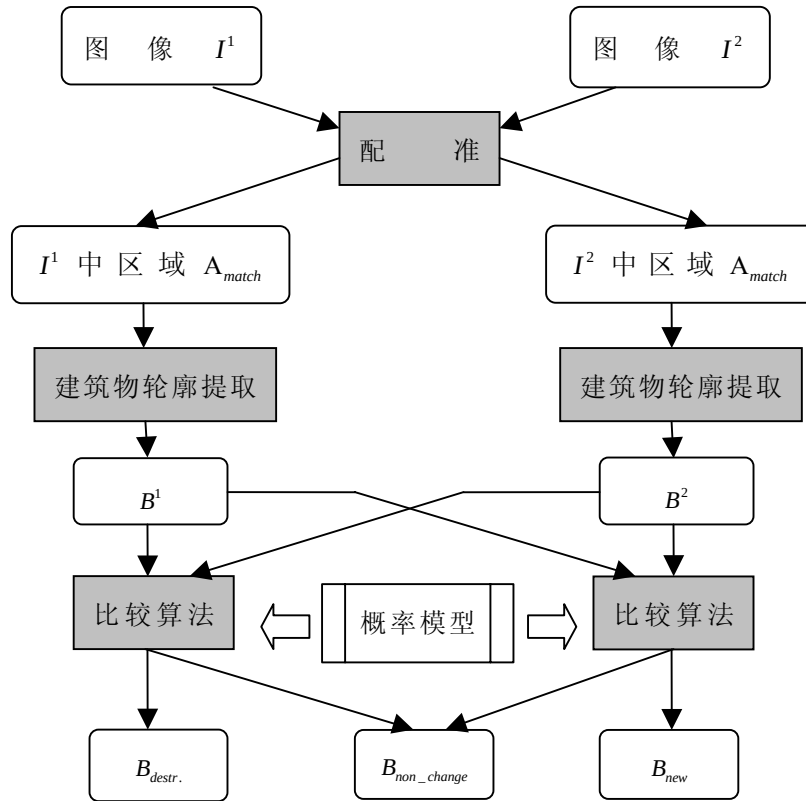


图 4-1 算法流程图

- ① 给定 $\Lambda_i^1, i \in S_1$ ，若存在 $\Lambda_j^2, j \in S_2$ ，满足 $\Lambda_i^1 = \Lambda_j^2$ ，则 $b_i^1 = b_j^2 \in B_{non_change}$

② 给定 $\Lambda_i^1, i \in S_1$ ，若不存在 $\Lambda_j^2, j \in S_2$ ，满足 $\Lambda_i^1 = \Lambda_j^2$ ，则 $b_i^1 \in B_{destr.}$

③ 给定 $\Lambda_j^2, j \in S_2$ ，若不存在 $\Lambda_i^1, i \in S_1$ ，满足 $\Lambda_i^1 = \Lambda_j^2$ ，则 $b_j^1 \in B_{new}$

其中 $\Lambda_i^1 = \Lambda_j^2 \Leftrightarrow \Lambda_i^1 \cap \Lambda_j^2 = \text{IDENTITY}$ 。在我们的算法中，通过一个概率模型表示上述准则并用以衡量 B_1 和 B_2 中每个建筑物发生变化的概率值。通过这个概率值，我们可以分别对 I^1 和 I^2 进行比较操作并阈值，最终判断 I^1 和 I^2 中建筑物的变化情况。图 4-1 反映了整个算法的流程。

4.3 配准

若要对两幅图像进行比较，我们首先必须确保图像能够正确的叠合在一起，也就是需要对它们进行配准操作。根据第一章对图像数据的介绍，我们所使用的数据都是经过地理编码的，也就是说图像 I^1 和 I^2 都已经相对于地理坐标系进行了矫正，我们可以根据图像像素点对应的地理坐标对它们进行配准。另外，由 4.5 节的描述将会知道本算法对配准精度的要求并不高，只要保证 I^1 和 I^2 中对应建筑物大部分区域都重合在一起即可，而采用地理编码进行配准足够保证这一点。

4.4 模型建立

假设图像 I^1 和 I^2 中的某个相同区域 i ，其轮廓为 C_i ，根据第三章中的概率模型，我们可以分别计算在 I^1 和 I^2 中 C_i 属于建筑物类的概率 $p_i^1 = p_i^1(x_i=1)$ 和 $p_i^2 = p_i^2(x_i=1)$ 。我们定义：

$$p_{non_change} = p_i^1 p_i^2 \quad \text{且 } t_1 < t_2 \quad \text{公式 4-1}$$

由上式，可以依据概率值的大小找出那些属于建筑物类且没有发生变化的轮廓，比如 $p_i^1=1$ 且 $p_i^2=1$ ，则 $p_{non_change}=1$ 。建筑物的变化主要包括建筑物的消失和出现，因此我们有： $p_{change} = p_{new} + p_{destr.}$ 。由于我们关心的是建筑物的变化情况，因此我们直接定义 p_{change} 如下所示：

$$p_{change} = p_i^2(1 - p_i^1) + p_i^1(1 - p_i^2) \quad \text{且 } t_1 < t_2 \quad \text{公式 4-2}$$

由上式，当 C_i 在 I^1 中是建筑物轮廓，在 I^2 中是非建筑物轮廓，即 I^1 中建筑物消失

时, 我们可以获得一个较大的 p_{change} 概率值, 比如 $p_i^1 = 1$ 且 $p_i^2 = 0$, 则 $p_{change} = 1$; 同样, 当 C_i 在 I^2 中是建筑物轮廓, 在 I^1 中是非建筑物轮廓, 即 I^2 中出现新建筑物时, 我们可以获得一个较大的 p_{change} 概率值, 比如 $p_i^1 = 0$ 且 $p_i^2 = 1$, 则 $p_{change} = 1$; 在建筑物没有发生变化时, 通常会获得较小的 p_{change} , 比如 $p_i^1 = 1$ 且 $p_i^2 = 1$, 则 $p_{change} = 0$ 。由以上分析, 我们可以通过公式 4-2 计算图像 I^1 和 I^2 中对应区域的轮廓变化概率值, 通过概率值的大小找到那些发生变化且属于建筑物类的轮廓: 若 $p_{change} > 0.5$, 则认为 C_i 发生了变化; 反之则认为 C_i 没有发生变化。若 C_i 发生了变化, 则进一步判断 p_i^1 和 p_i^2 的大小: 若 $p_i^1 > p_i^2$, 则建筑物消失; 反之则新建筑物出现。

4.5 比较算法

比较操作的目的是找到图像 I^1 和 I^2 中对应的轮廓进行比较判断它们是否发生变化。

在进行比较操作之前, 首先对图像 I^1 和 I^2 应用第三章的建筑物轮廓识别算法, 需要注意的是, 在识别算法最后, 我们并不进行建筑物的选择, 也就是说保留融合算法的结果。这样做的原因是 4.4 节中的轮廓变化模型本身具有区分建筑物轮廓和非建筑物轮廓的作用。

在实际应用中, 公式 4-2 介绍的变化检测模型, 主要具有两方面的问题: 一是无法保证 S_1 和 S_2 中元素的一一对应关系; 二是 $\Lambda_i^1 = \Lambda_j^2$ 很难被严格满足。因此经常会造成这种情况的发生: $b_i^1, i \in S_1$ 无法找到一个 $b_j^2, j \in S_2$, 能够保证 $\Lambda_i^1 = \Lambda_j^2$ 。为了解决这个问题, 我们转而考虑那些具有一定相交区域的轮廓 $b_i^1, i \in S_1$ 和 $b_j^2, j \in S_2$, 且 $\Lambda_i^1 \cap \Lambda_j^2 > THRESHOLD$, 其中 $THRESHOLD$ 是最小相交面积阈值, 其作用是忽略那些相交面积很小的情况。根据轮廓对应的 $p_i^1(x_i = 1)$ 、 $p_j^2(x_j = 1)$ 和公式 4-2 可以计算 b_i^1 的变化概率 p_{change} 的大小。若 $b_i^1, i \in S_1$ 同时和几个 $b_j^2, j \in S_2$ 相交, 则计算每对相交物体的 p_{change} , 把最大的 p_{change} 作为物体 b_i^1 的最终 p_{change} 。若 S_2 中没有和 b_i^1 相交的轮廓, 则 $p_{change}^i = p_i^1$ 。反复使用上述算法, 直至遍历整个 $B_1 = \{b_i^1\}, i \in S_1$ 得到所有 b_i^1 的变化概率值, 最后通过阈值删除那些没有变化的建筑物轮廓, 得到图像 I^1 中变化的建筑物轮廓集——即 $B_{destr.}$ 。按同样的方法遍历 $B_2 = \{b_j^2\}, j \in S_2$ 中的元素可以获得图像 I^2 中变化的建筑物轮廓集——即 B_{new} 。

使用比较算法计算 $B_{destr.}$ 的伪代码如下所示:


```

procedure COMPARE_PROCESS( $B_1, B_2, B_{destr.}$ )
 $B_{destr.} \leftarrow \emptyset$ 
for  $i = 1$  to  $\dim(S_1)$ 
     $\max\_probability = 0$ 
    for  $j = 1$  to  $\dim(S_2)$ 
         $\text{inter\_area} = \text{area}(\Lambda_i^1 \cap \Lambda_j^2)$ 
         $p_{change}^{temp} = p_j^2(1 - p_i^1) + p_i^1(1 - p_j^2)$ 
        if  $\text{inter\_area} > \text{THRESHOLD} \ \&\& \ p_{change}^{temp} > \max\_probability$ 
             $\max\_probability = p_{change}^{temp}$ 
        endif
    repeat
        if  $\max\_probability > 0$ 
             $p_{change}^i = \max\_probability$ 
        else
             $p_{change}^i = p_i^1$ 
        endif
    repeat
for  $i = 1$  to  $\dim(S^1)$ 
    if  $p_{change}^i > 0.5$ 
         $B_{destr.} = B_{destr.} \cup b_i^1$ 
    endif
repeat
end COMPARE_PROCESS

```

4.6 实验结果及分析

本章介绍的变化检测算法是基于第三章的建筑物轮廓识别算法，且使用的数据也类似，我们使用和第三章算法相同的参数设置。所不同的是由于图像来自于两个不同时相，成像条件不同，因此需要分别设置两幅图像的太阳光方位

角。在比较算法中，设置两轮廓相交面积的阈值 THRESHOLD 为轮廓自身面积的 50%。

识别算法的参数来自于对 2002 年建筑物训练集的参数优化而没有包括 2003 年的训练集，之所以不使用 2003 年的训练集是因为在实际中，新得到的数据往往是用于更新旧 GIS 以获得新的 GIS 的，若要使用新数据作训练集，则需要新的 GIS 或者人工标记，这都是和实现自动化相悖的。

首先使用合成图像做简单试验进行算法验证。原始图像来自 2002 年北京 QUICKBIRD 卫星图像，经过对图像上某些地区覆盖简单的纹理合成新的图像作为假想的不同时相图像。变化检测算法如图 4-1 所示，图中同时展示了经过识别算法后未阈值的轮廓。从图中可以看出，在使用合成图像即不考虑图像间灰度分布及场景区别的情况下，变化检测结果良好，所有变化的建筑物都能够被很好的识别出来，这也验证了我们提出的变化检测模型的有效性。

随后，使用变化检测算法于 2002 年和 2003 年的真实图像上，图 4-2 和图 4-3 是两个实例。通过上一章的分析知道建筑物轮廓识别算法对于很多具有较复杂纹理的建筑物的识别率不高，因此在选择变化检测测试集时，尽量避免选择过于复杂的建筑物。这里，总共使用了 8 幅图像包含大约 62 个变化建筑物进行测试，并对变化建筑物的检测正确率进行了统计，结果如表 4-1 所示，其中用到的参考数据（ref 角标）是通过人工判别得到的。

可以看出，我们的算法能够识别出大部分变化的建筑物。但是，相对于建筑物轮廓识别算法的识别率来说，建筑物的变化检测率是比较低的，这是因为我们的建筑物变化检测算法是基于建筑物轮廓识别算法，并同时应用于两幅图像上的，两幅图像的总建筑物识别率是每幅图像识别率的乘积。另外一个重要的原因是两个不同时间图像的场景及灰度分布是不同的，造成使用 2002 年训练集图像优化出的参数在 2003 年图像上进行建筑物识别无法达到最优的效果，从而造成总体检测率的下降。其次，在进行变化检测的比较算法时，对于某轮廓同时和另一图像中多个轮廓相交的问题——即无法达到一一映射的问题，目前所使用的解决方法还不完善。实际上造成这种多重映射的情况很复杂，但大多数情况下都是由于识别算法的结果不够准确造成的。一种典型的情况是识别算法在一幅图像上完整的提取出了某个建筑物；而在另一幅图像上这个建筑物却没有被完整的提取——比如被分成了好几个小部分。这样，在进行变化检测

的比较算法时，那个完整的建筑物轮廓会和另一幅图像上这个建筑物各个部分的轮廓进行多次比较，这样得到的变化概率显然很难正确反映建筑物的实际变化情况。

image	NUM_{change}^{ref}	$NUM_{change}^{dec_right}$	$NUM_{change}^{dec_error}$	$DecR_{change}$	$CorR_{change}$
1	6	4	3	0.67	0.57
2	9	6	0	0.67	1.00
3	8	5	5	0.63	0.50
4	8	6	4	0.75	0.6
5	5	4	2	0.8	0.67
6	9	6	4	0.67	0.6
7	10	9	0	0.90	1.00
8	7	5	4	0.71	0.56
Total	62	45	22	0.72	0.67

表 4-1 NUM_{change}^{ref} ：真实变化的建筑物数目； $NUM_{change}^{dec_right}$ ：监测到正确的变化建筑物数目； $NUM_{change}^{dec_error}$ ：监测到错误的变化建筑物数目； $DecR_{change}$ ：检测率， $DecR_{change} = NUM_{change}^{dec_right} / NUM_{change}^{ref}$ ； $CorR_{change}$ ：正确检测率， $CorR_{change} = NUM_{change}^{dec_right} / (NUM_{change}^{dec_right} + NUM_{change}^{dec_error})$



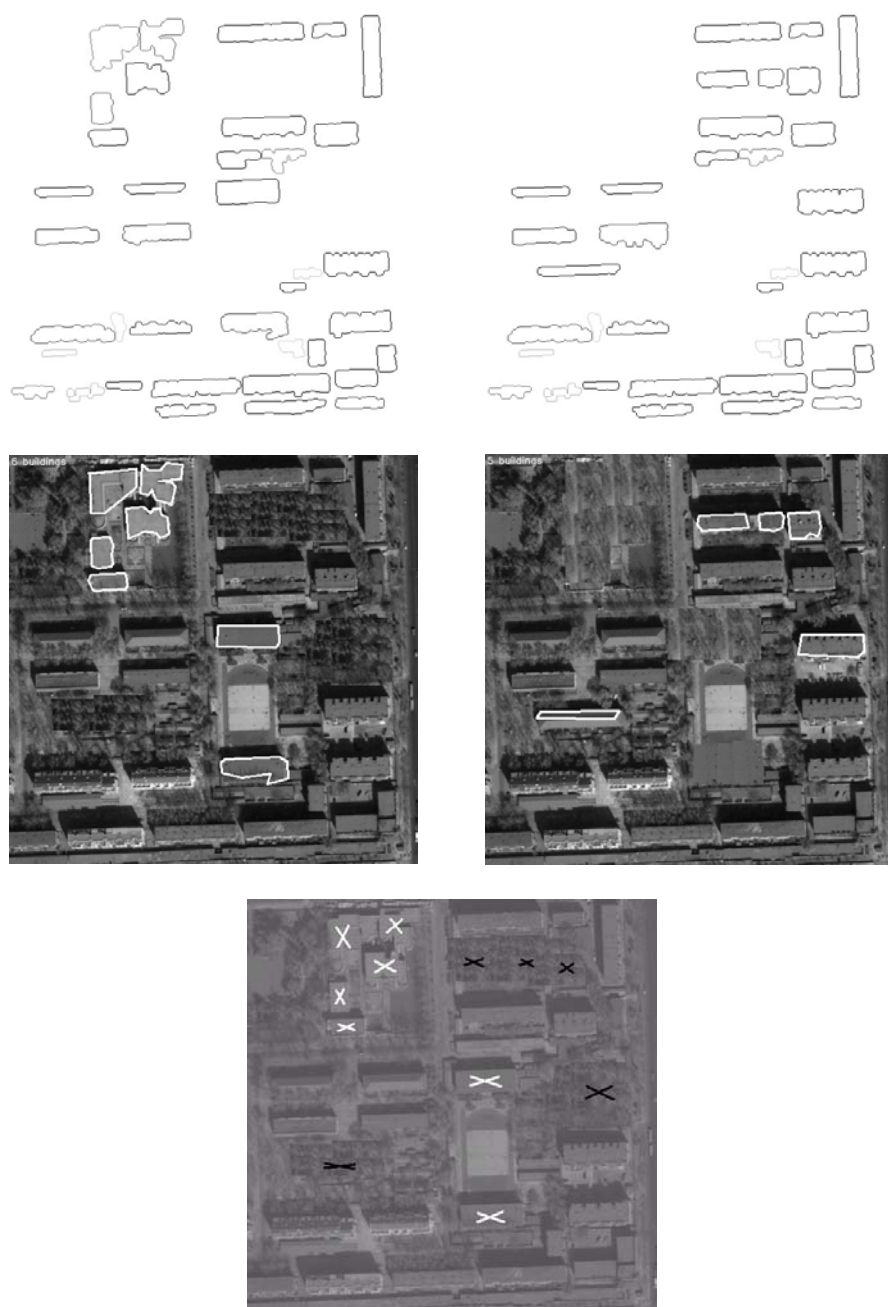
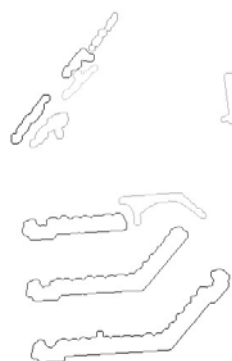


图 4-1 合成图像变化检测。第一排分别是同一地区“旧”图像和“新”图像；第二排是没有经过阈值的建筑物轮廓识别算法结果，图像中的灰度值是轮廓属于建筑物轮廓类别的概率，具体含义可参见第三章；第三排分别是在新旧两幅图像上找到的发生变化的建筑物；第四排图像是人工标记结果，其中的小叉表示所对应建筑物重心及长短轴：白叉表示消失的建筑物，黑叉表示新出现的建筑物。



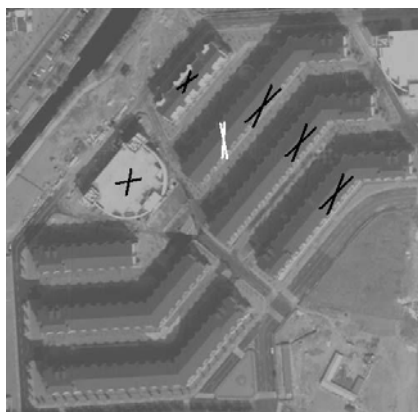


图 4-2 表 4-1 第 2 号真实图像变化检测。第一排分别是同一地区 2002 年图像和 2003 年图像；图像意义见图 4-1 注释。

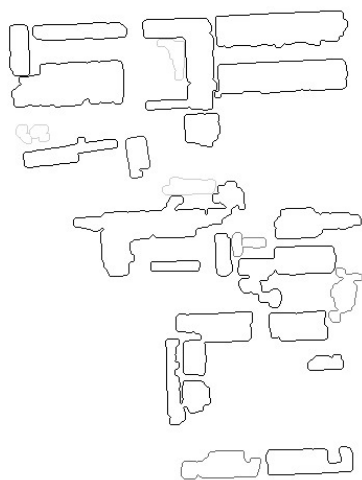




图 4-3 表 4-1 第 7 号真实图像变化检测。图像意义见图 4-2 注释。

4. 7 小结

在这一章中，我们提出了一种相对简单的，基于概率模型的建筑物变化检测算法，这种算法的主要步骤是首先识别出两幅图像中的建筑物区域，然后结合预先定义的建筑物变化概率模型，运用比较算法计算两幅图像中相交轮廓的变化概率并通过阈值得到发生变化的建筑物轮廓。通过实例可以看到这种算法对于一般纹理复杂度的建筑物的变化具有较好的检测能力。但是，我们也看到由于这种算法是基于建筑物轮廓识别算法的，因此建筑物轮廓识别算法在很大程度上决定了变化检测算法的效果。在以后的工作中，我们可以通过提高建筑物识别算法的正确率提高变化检测算法的正确率，另外，还可以通过比较两个

不同时相图像的直方图消除它们之间的差别，使得应用旧图像上训练得到的优化参数在新图像上进行建筑物识别能获得高的正确率。

第五章 基于 DRF 模型的结构物体变化检测算法

5.1 引言

在这一章，将介绍一种寻找结构物体变化区域的算法。所谓结构物体，即具有特定结构特征的物体，本章主要关注建筑物。一般建筑物都具有直线边缘、直角、多边形等特定的结构特征。算法的主要思想是基于 DRF (Discriminative Random Field) 模型，结合图像中格子这种基本“元素”之间的关系衡量它们的结构变化性。我们扩展了已有的 DRF 模型并用于结构物体的变化检测，在假想数据和真实数据上获得了初步的实验结果。

下面，将首先介绍算法流程，然后分别对流程中各个步骤作以详细描述，最后，将给出试验结果并进行分析。

5.2 算法流程

类似于第四章，对于不同时间拍摄的同一地区的图像 I^1 和 I^2 ，首先需要对它们进行配准；然后在每幅图像上分割出规则大小的格子作为算法的基本“元素”，计算反映图像 I^1 和 I^2 中对应元素结构特性变化的特征；基于变化检测 DRF 模型，利用推断算法判断每个元素是否发生变化，从而得到最终结果。类似于第三章中的建筑物识别概率模型，我们需要首先利用先验知识（训练集）对模型参数进行优化。图 5-1 是整个算法的流程图。这里的配准算法类似于第四章中的介绍，因此不再赘述。

5.3 模型建立

5.3.1 DRF 模型回顾

DRF 模型 (Discriminative Random Field) 是由 Kumar 在 2003 年提出的，主要用于自然场景中建筑物的识别。这种模型是建立在规则格子这种基本元素上的，通过按照预定规则格子的大小 $n_1 \times m_1$ ，把图像分割成相同格子的区域(如图 5-2 所示)。假设所有格子构成图结构 $G = (S, E)$ ，其所有节点集合为 S ， E

是图结构中边的集合。

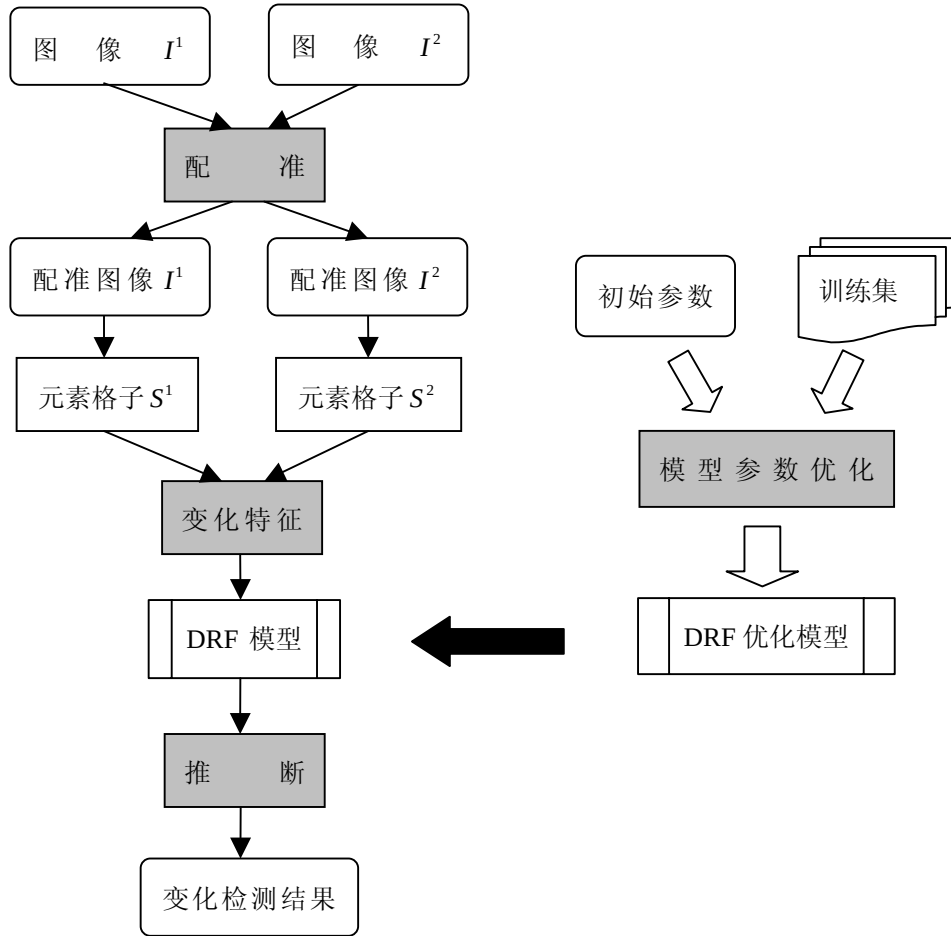


图 5-1 算法流程图

且有： $\dim(G) = (n/n_1)/(m/m_1)$ ，其中 $n \times m$ 代表图像的尺寸。S 中每个节点对应的观测值表示为： $\mathbf{y} = \{\mathbf{y}_i\}_{i \in S}$ ， $\mathbf{y}_i$ 表示第 i 个节点对应的数据或者特征，Kumar 使用 DRF 进行结构物体识别，选择 $\mathbf{y}_i$ 为反映格子区域内结构特性的特征向量。S 中节点相应的标记为 $\mathbf{x} = \{\mathbf{x}_i\}_{i \in S}$ ，对图像进行二元分类，定义 $x_i \in \{-1, 1\}$ ，其中 1 和 -1 分别对应于结构物体标记与非结构物体标记。由 Gibbs 分布[57]，随机变量 $\mathbf{x}$ 和 $\mathbf{y}$ 的概率分布可表示为：

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{-U} \quad \text{公式 5-1}$$

其中 Z 为归一化函数，依据 Gibbs 分布和 markov 场的一致性[53, 55]，势能函数 U 可以表示为：

$$U = -(\sum_{i \in S} A_i(x_i, \mathbf{y}) + \sum_{i \in S} \sum_{N_i} \beta_{ij} I_{ij}(x_i, x_j, \mathbf{y})) > 0 \quad \text{公式 5-2}$$

其中 A 为联合势，反映单独格子区域内的信息；交互势 I 反映格子 i 与其邻域 N_i 中格子的相互关系，也是空间连续性的度量——这是基于一般物体的各个部分在空间上是连续的实事，也是 Markov 场的基本性质。通过定义相应的联合势函数、交互势函数及邻域系统，可以完全定义后验概率分布 $P(\mathbf{x}|\mathbf{y})$ 。这样，寻找建筑物格子区域的问题转化为寻找后验概率 $P(\mathbf{x}|\mathbf{y})$ 最大值的问题：当 $P(\mathbf{x}|\mathbf{y})$ 达到最大值时，其对应标记配置 $\mathbf{x}$ 就是最优的分类标记配置。图 5-2 是使用 DRF 对一般场景中建筑物的识别结果，其中高亮格子为算法找到的建筑物格子，其它则为非建筑物格子。

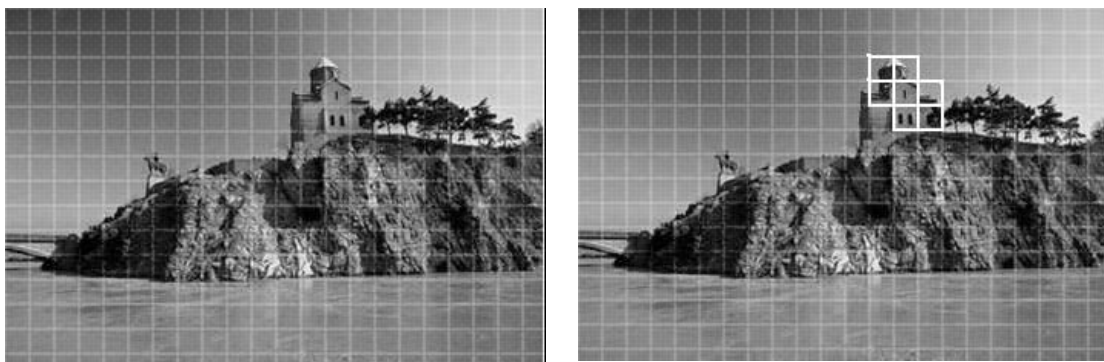


图 5-2 自然场景建筑物识别。图像大小：350x230。格子大小为：16x16。

5.3.2 变化检测 DRF 模型

以上介绍的 DRF 模型主要运用于二维图像中，变化检测则需要考虑不同时相的图像间的变化，相当于增加了一个新的维度——时间，因此需要对已有的 DRF 模型进行扩展。变化检测 DRF 模型的主要思想是同时考虑格子在时域上的不连续性和在空间上的连续性。

类似于 DRF 模型，我们依然采用规则格子作为图像 I^1 和 I^2 的基本元素（两幅图像已经过了配准）。不同的是，这里二元分类转化为寻找变化的结构物体和没有变化的结构物体： $x_i \in \{-1, 1\}$ ，其中 -1 和 1 分别对应于没有变化的结构物体标记和变化的结构物体标记。由标记的定义，我们需要寻找能够表示结构变化的特征。假设表示 I^1 和 I^2 的变化的特征可以表示为： $\mathbf{y} = \{\mathbf{y}_i\}_{i \in S}$ ，其中 $\mathbf{y}_i$ 为两幅图像对应位置 i 的结构变化特征。沿用 DRF 模型中的符号，由公式 5-1 与公式 5-2，

后验概率分布为：

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z} e^{\sum_{i \in S} A_i(x_i, \mathbf{y}) + \sum_{i \in S} \sum_{N_i} \beta_{ij} I_{ij}(x_i, x_j, \mathbf{y})} \quad \text{公式 5-3}$$

其中联合势 A 主要反映元素在时域上的不连续性，即图像 I^1 和 I^2 中位置 i 的结构特性变化情况，具体定义为：

$$A_i(x_i, \mathbf{y}_i) = \log(\sigma(x_i \omega^T h_i(\mathbf{y}_i))) \quad \text{公式 5-4}$$

其中 ω^T 是模型参数，函数 $h(\bullet)$ 是非线性转换函数，作用和意义同第三章建筑物轮廓识别中的转换函数 $f(\bullet)$ ；函数 $\sigma(\bullet)$ 为逻辑函数（定义见第三章）。取对数的意义在于不考虑交互势的情况下，整个 DRF 模型将退化成简单的逻辑分类模型。

交互势 I 主要反映位置 i 和其邻域 $j \in N_i$ 的相似性，即空间上的连续性。考虑这种连续性是因为若某物体的一部分发生改变则此物体其它部分也会发生改变。联合势的具体定义为：

$$I_{ij}(x_i, x_j, \mathbf{y}) = K x_i x_j + (1 - K) (2 \sigma(x_i x_j \mathbf{v}^T u_i(g(\mathbf{y}_i, \mathbf{y}_j))) - 1) \quad \text{公式 5-5}$$

其中 $\mathbf{v}^T$ ， K ， β_{ij} 为模型参数， K 的取值范围是 $[-1, 1)$ 。 $u(\bullet)$ 是非线性转换函数， $g(\bullet)$ 用于计算向量 $\mathbf{y}_i$ 和 $\mathbf{y}_j$ 之间的差别向量，算法中我们计算向量间的绝对距离。联合势中第一项是基本的 Ising 模型[53]，表示空间上的连续性；但是仅仅使用 Ising 模型容易导致过平滑[53]，因此引入了一个附加的 DA (Discontinuity adaptive) 模型[53]，它是一个数据依赖项，其作用是通过考察数据（特征）来缓解或者修正 Ising 模型带来的过平滑。 K 的作用是调整 Ising 模型和 DA 模型之间的比重关系。 β_{ij} 的作用在于调节联合势的整体平滑性。对于各向同性的邻域系统， β_{ij} 为常数；对于各向异性的邻域系统， β_{ij} 根据方向性不同取不同常数。在模型的实际应用中，我们定义邻域系统为四邻域且仅考虑成对元素间的相互关系。另外，假设图结构 G 是各向同性的，则对于邻域系统 N_i ， β_{ij} 取常数 β 。

5.4 参数估计

类似于第三章，为了让模型具有正确的分类能力，需要首先利用先验知识（训练集）对模型参数进行优化以获得最优参数。DRF 模型参数集合是： $\theta = \{\omega, \mathbf{v}, K, \beta\}$ 。同一般 MRF 模型参数优化比较，由于 DRF 模型直接表示后验概率分布，因此不需要像 MRF 模型那样分别估计模型中的参数并且需要做参数独立性

假设，所有模型参数可以同时进行优化。MRF 模型中的大部分参数优化方法都可以很容易的应用于 DRF 模型中，在这里，我们采用最大似然方法优化 DRF 模型参数。参数估计中包括对归一化函数 Z 的估计，这是一个 NP 困难问题，一般可以采用一些近似算法——比如平均场，伪似然函数[53]等方法。这里采用相对简单有效的伪似然函数进行参数估计，有：

$$\begin{aligned}\hat{\theta}^{ML} &= \arg \max_{\theta} \prod_{m=1}^M \prod_{i \in S} P(x_i^m | \mathbf{x}_{N_i}^m, \mathbf{y}^m, \theta) \\ \theta &= \{ \boldsymbol{\omega}, \mathbf{v}, K, \beta \} \\ K &\in [-1, 1)\end{aligned}\tag{公式 5-6}$$

其中 M 表示训练集中图像的数目， m 是其中图像的标号。生成训练集的方法类似于第三章中介绍的方法。另外，每个位置 i 对应的概率分布可以表示为：

$$\begin{aligned}P(x_i | \mathbf{x}_{N_i}, \mathbf{y}, \theta) &= \frac{1}{Z_i} \exp(A(x_i, \mathbf{y}) + \sum_{j \in N_i} I(x_i, x_j, \mathbf{y})) \\ Z_i &= \sum_{x_i \in \{-1, 1\}} \exp\{A(x_i, \mathbf{y}) + \sum_{j \in N_i} I(x_i, x_j, \mathbf{y})\}\end{aligned}\tag{公式 5-7}$$

公式 5-6 中的伪似然函数可以通过带约束的线性搜索方法进行最大化[46]。由于伪似然函数不是一个凸函数，所以为了避免优化算法陷入局部最小值，需要较好的参数初始值。可以看到，联合势与交互势中的数据依赖项（DA 模型）都是由逻辑函数构成的，由第三章的证明可知，它们都是凸函数，因此可以首先分别对这两个凸函数中的参数 $\boldsymbol{\omega}$ 和 $\mathbf{v}$ 进行优化（方法见第三章），随后，把优化后的参数作为整个 DRF 模型参数优化算法的初始参数。采用这种两步优化的参数优化方法，可以有效避免陷入局部极值从而得到正确的 DRF 模型优化参数。

5.5 结果推断

当给定两幅图像 I^1 和 I^2 时，我们的目标是寻找图像中格子对应标记集合 $\mathbf{x} = \{x_i\}_{i \in S}$ 的最优配置。变化检测 DRF 模型能够衡量某个元素是否发生变化，当整个图像的 DRF 模型达到最大值时，所对应的 $\mathbf{x}$ 便是最优配置标记。寻找最优配置 $\mathbf{x}$ 是一个组合问题，需要搜索集合 $\mathbf{x}$ 中各个元素可能取值的各种组合方式。因此，结果推断的过程也是一个参数优化的过程，同模型参数优化不同的是，结果推断面对的是离散组合空间的优化问题。一般来说，我们可以使用 MAP (Maximum A Posteriori)、MPM (Maximum Posterior Marginal)、ICM (Iterated Conditional

Modes) 等最大化算法进行优化。经过证明[54, 55], MAP 算法应用于 Ising 模型时, 若交互势参数 β 取值较大时, 往往很难获得满意的优化结果; 另外, 对于 MPM, 由于过大的计算量往往成为 NP 困难问题。因此, 我们选择一种相对简单的 ICM 最优化方法[56]。ICM 算法的基本思想是反复最大化局部条件概率函数, 直到算法收敛。如下式所示:

$$x_i \leftarrow \arg \max_{x_i} P(x_i | \mathbf{x}_{N_i}, \mathbf{y}) \quad \text{公式 5-8}$$

ICM 往往能找到后验概率的局部的极值并被证明即使在 β 取值过大导致 MAP 效果太差的情况下也能获得较为满意的结果。但是, ICM 的关键问题是需要较好的初始标记配置 $\mathbf{x}_{init}$ 才能获得较好的优化结果。为了获得较好的初始标记, 我们舍弃 DRF 模型中的交互势, 仅使用联合势函数, 由前面的分析得知, 此时后验概率模型退化为简单的逻辑分类模型:

$$P(x_i | \mathbf{x}_{N_i}, \mathbf{y}) = \sigma(x_i \boldsymbol{\omega}^T \mathbf{h}_i(\mathbf{y})) \quad \text{公式 5-9}$$

然后采用标准 MAP 优化方法最大化公式 5-9 便能获得 ICM 的初始输入标记配置

$\mathbf{x}_{init}$ 。

5.6 特征选择

本章主要处理的是建筑物的变化检测问题, 因此对应的结构特征 $\mathbf{y}$ 相应是能够反映建筑物变化的结构特征。我们主要考虑具有多边形轮廓的建筑物, 这样的建筑物主要具有如下几方面特征: 轮廓主要由直线构成, 直线间互相平行或者成直角等。对应的建筑物变化的特征则是: 轮廓直线方向的变化, 直线间相互关系的变化, 直线的消失或者出现等。

类似于第三章中轮廓形状特征的计算方法, 我们分别计算图像 I^1 和 I^2 中位置 i 的梯度-方向直方图 (方法参见第三章), 这样, 我们能获得位置 i 在图像 I^1 和 I^2 中的两幅直方图: HISTO^1 和 HISTO^2 。分别计算两幅直方图中对应的二阶矩, 分别表示为: MOMENT_2^1 和 MOMENT_2^2 。计算二阶矩特征的目的是为了捕捉直方图中的“尖峰性”。一般来说, 若位置 i 对应的格子区域中存在明显的直线——即格子内可能包括建筑物的部分轮廓区域, 则在其梯度-角度直方图中表现为具有很明显的尖峰; 反之, 若格子所属的区域是非建筑物结构区域——比如植被, 则直方图

趋于均匀，不会出现明显的尖峰。因此，若建筑物发生了变化——比如出现新建筑物或者旧建筑物消失，则在图像 I^1 和 I^2 的某些位置中这些矩的大小也会有较明显的变化，所以，考察梯度-角度直方图的矩变化可以衡量是否有建筑物发生变化，向量 $\mathbf{y}_i$ 中的第一个特征为：

$$\text{diff}_{\text{moment}2} = |\text{MOMENT}_2^1 - \text{MOMENT}_2^2| \quad \text{公式 5-10}$$

前面提到，建筑物结构特征中一个重要的特征就是轮廓多边形直线段间具有平行性或者垂直性，因此，若不同时间同一区域内直线的平行性或者垂直性发生了明显的变化，则表明该区域内可能有建筑物发生了变化。我们分别计算 HISTO^1 和 HISTO^2 中第一大和第二大梯度大小峰值对应的梯度角度： $\delta_{1st}^1, \delta_{2nd}^1$ 和 $\delta_{1st}^2, \delta_{2nd}^2$ ，则特征 $|\cos 2(\delta_{1st}^1 - \delta_{2nd}^1)|$ 和 $|\cos 2(\delta_{1st}^2 - \delta_{2nd}^2)|$ 分别反映位置 i 在图像 I^1 和 I^2 对应区域内直线段的平行性或者垂直性。因此，它们之间的差值可以反映直线段的平行性或者垂直性变化的情况。定义向量 $\mathbf{y}_i$ 第二个特征为：

$$\text{diff} = |\cos 2(\delta_{1st}^1 - \delta_{2nd}^1) - \cos 2(\delta_{1st}^2 - \delta_{2nd}^2)| \quad \text{公式 5-11}$$

另外，若建筑物发生变化，则其多边形轮廓直线的方向可能发生变化，在梯度-角度直方图上反映为梯度尖峰对应的梯度角度发生变化，因此，这些角度的变化可以反映建筑物结构特性的变化。则向量 $\mathbf{y}_i$ 第三个特征可以表示为：

$$\text{diff}_{\text{angle}} = |\cos(\delta_{1st}^1 - \delta_{1st}^2)| \quad \text{公式 5-12}$$

以上计算的特征都是在位置 i 区域内的特征，也就是特定格子大小区域内的特征。为了增加鲁棒性，可以考虑位置 i 周围区域对其的影响，因此我们引入了多尺度特征：即扩展位置 i 的区域面积，然后在扩展后的区域内计算上述变化特征，并把这些特征作为位置 i 的一部分特征加入到特征向量 $\mathbf{y}_i$ 中。在算法的具体应用中，我们采用三级多尺度特征，即在预先定义的规则格子原始尺寸的基础上，扩大两次格子尺寸，每次扩大的尺寸是上一次尺寸的 2 倍。假设格子尺寸为 $n_1 \times m_1$ ，则两次扩大后的格子尺寸分别为： $2n_1 \times 2m_1$ ， $4n_1 \times 4m_1$ 。

5.7 实验结果及分析

这一小结中，变化检测 DRF 模型将被应用于“玩具”图像、合成图像和真实图像中的结构物体变化检测中，我们将对相应的实验结果及存在的问题作以

简单的分析。

5.7.1 玩具图像结构变化检测试验

所谓玩具图像，即完全由人工绘制的简单图像。在玩具图像上试验，是为了检测我们的变化检测模型的基本能力和性质。我们使用了 6 对具有结构变化的玩具图像进行参数训练，然后利用优化参数对新的玩具图像进行变化检测。为了检测模型各方面的能力，在检测图像中包括了直线、曲线、规则多边形、不规则闭合轮廓、规则纹理和不规则纹理等内容及其相应的变化。同时我们还对不同格子尺寸对变化检测结果的影响进行了实验。图 5-3 例示一个实验结果。

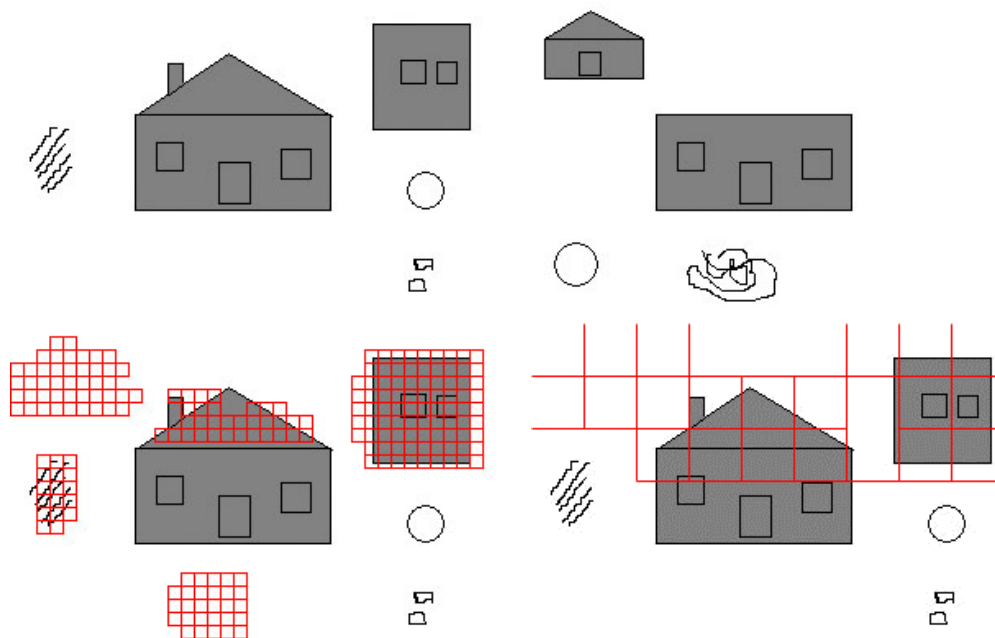


图 5-3 玩具图像试验。上排是两幅具有结构变化的玩具图像；左下是使用 8x8 大小格子的检测结果和第一幅图像的叠加。右下是使用 32x32 大小格子的检测结果在第二幅图像的叠加。

从上图可以看出，两种大小的格子都能够较好的检测出图像中大部分结构变化区域，同时，图像中的圆形轮廓的变化都没有被检测出来，这证明我们的特征选择能够正确反映结构变化特性，而且我们的变化检测 DRF 模型能够正确利用这些特征进行正确分类。我们也可以看到第一幅图像左边的那些直线纹理的变化在采用 8x8 格子大小时被检测出来，而在采用 32x32 格子大小时则没有

被检测出来。虽然不管采用小尺寸格子还是采用大尺寸格子都能在纹理覆盖的区域得到明确的结构变化特征，但是由于 DRF 模型具有交互势，还会考察每个位置和它周围位置之间的连续性（平滑性）：对于小尺寸的 8x8 格子，由于纹理区域覆盖了好几个格子，每个格子对应的位置都发生了结构变化，它们之间存在相似性，在模型中反映为具有较高的交互势，因此在结果推断时会同时被认定发生了结构变化；而对于大尺寸的 32x32 格子，纹理区域只被一个格子覆盖，其周围的格子不具有结构变化，在模型中虽然能够得到较大的联合势，但是由于交互势较小，因此在进行结果推断时，纹理对应的格子被认为没有发生结构变化。类似的，我们考察图像中下部中间区域曲线纹理区域的变化，由于在 8x8 格子中，每个格子只能覆盖曲线的一小部分，而这一小部分会表现为直线特征，因此最终曲线变化区域被不正确的认为是结构变化区域；而对于 32x32 格子，曲线区域在单独的格子中无法表现出直线特征，因此也就无法被认定为发生了结构变化。从以上分析中可以看出两点：第一，交互势在模型中发挥着重要的作用，在很多情况下能够通过交互势衡量元素间的平滑性，从而排除那些在空间上不连续的孤立元素；第二，格子尺寸的选择对计算结果具有一定的影响，在实际应用中，我们应该选择合适的格子尺寸以避免把非结构变化作为结构变化。

5.7.2 合成图像结构变化检测试验

我们选用来自 2002 年的 QUICKBIRD 卫星图像，经过对某些地区覆盖简单的纹理合成新的图像作为假想的不同时相图像。进行合成图像试验的目的是为真实图像的变化检测做准备，主要考察算法的有效性和选取合适的格子尺寸。我们使用 6 对合成图像包含大约 50 个变化建筑物作为训练集图像。根据格子尺寸的不同，模型参数优化时间从 3 分钟到 10 分钟，使用 ICM 进行结果推断时间从几秒钟到几十秒。图 5-4 例示了一些实验结果。

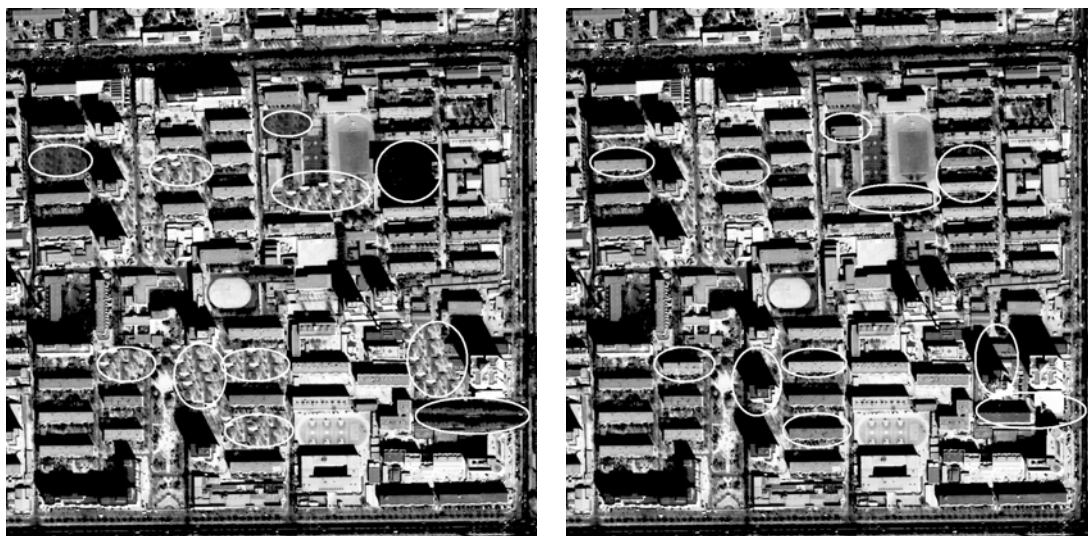
可以看出，对于来自于真实场景的合成图像，变化检测 DRF 模型仍然能够获得较好的效果。但是采用不同尺寸格子的实验结果也有所不同：较小尺寸的格子往往能够获得相对准确的变化建筑物区域，但是也可以看出很多情况下无法完全找到变化建筑物的所有部分；较大尺寸的格子，能够覆盖更多变化建筑物的区域，但是在每个格子中可能包括更多的冗余信息。由于变化检测 DRF 模型

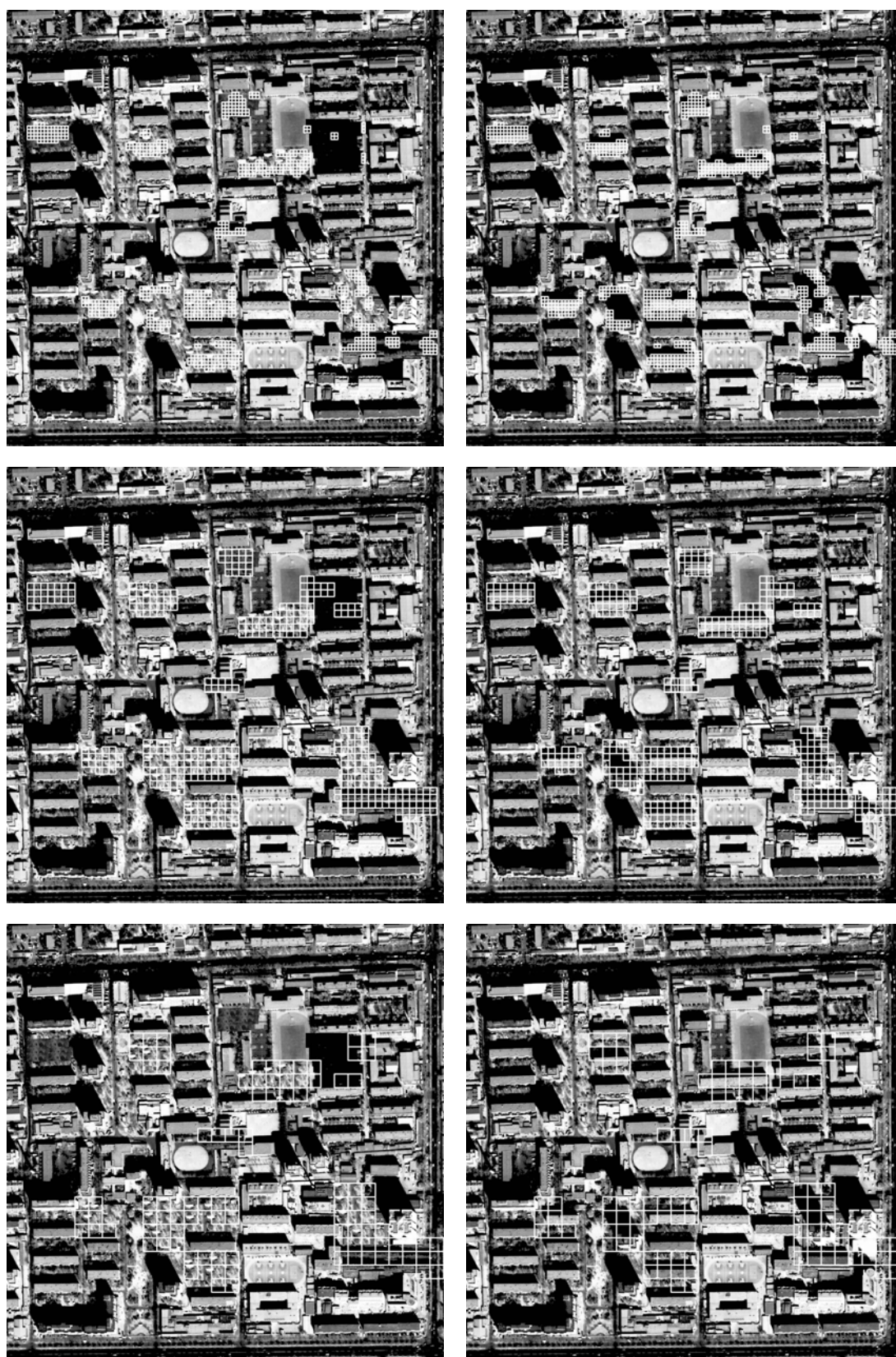
的目的仅仅是寻找变化结构物体的区域，在实际应用中，进行变化检测最终需要得到变化物体的准确轮廓。大尺寸的格子能够覆盖更多的变化物体区域，为进一步的轮廓识别提供更多更完整的信息，因此在真实图像的变化检测中，我们倾向于采用大尺寸格子作为基本元素。

5.7.3 真实图像结构变化检测试验

我们选用来自 2002 和 2003 年相同地区的 QUICKBIRD 卫星图像进行试验。我们使用 5 对图像包含大约 100 个变化建筑物作为训练集图像。格子尺寸选择 64x64，模型参数优化时间为 10 分钟，使用 ICM 进行结果推断时间仅几秒钟。从图 5-5 可以看出，大部分图像中发生变化的建筑物都能够用格子标记出来，证明变化检测 DRF 模型在某些真实图像上能够获得较好的结果。

但是也需要指明，在大量真实图像的实验中，只有少数图像能获得较好的结果，很多情况下的检测结果都不能让人满意，变化检测 DRF 模型的检测效果相对于玩具图像或者合成图像有了明显的下降，造成检测质量下降的原因主要是因为真实图像间场景的剧烈变化，成像条件、气候、植被等因素都会改变物体周围的环境，可能造成变形、遮挡等情况的发生。未发生变化的物体区域在一幅图像中很明显的结构特征在另一幅图像中可能并不明显；同时，发生变化的物体区域可能由于周围环境的影响在两幅图像中都表现出很强的结构特征。这样便会影响特征的分辨率，从而影响模型优化参数的质量，造成模型分类能力的下降。





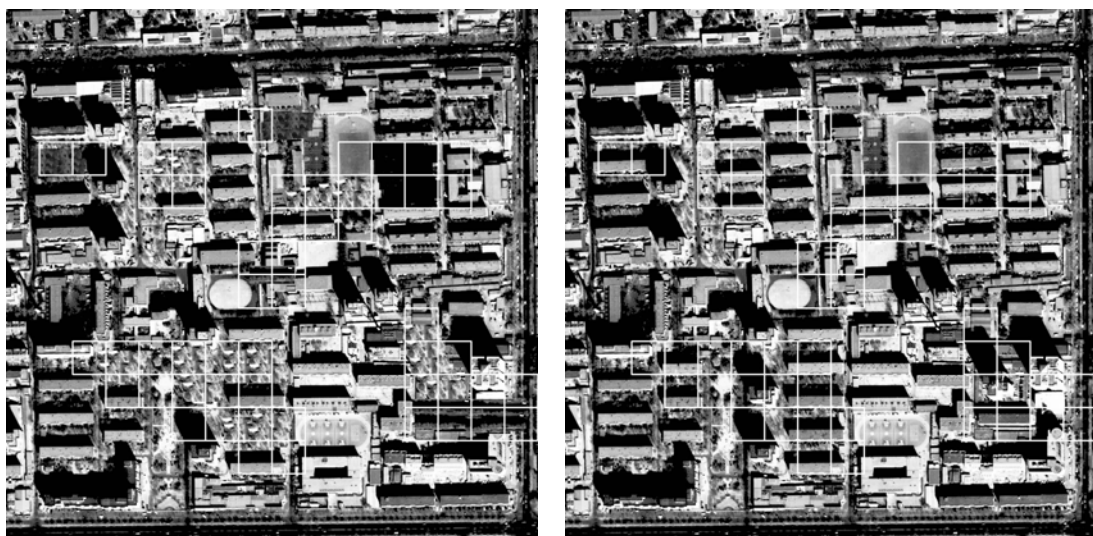


图 5-4 合成图像试验。图像大小：1024x1024。第一排是原始图像，人工标记出变化区域。从第二排图像开始，每排图像分别对应不同格子尺寸的变化检测结果在两幅合成图像上的叠加。从上到下依次格子尺寸依次是：8x8，16x16，32x32，64x64。

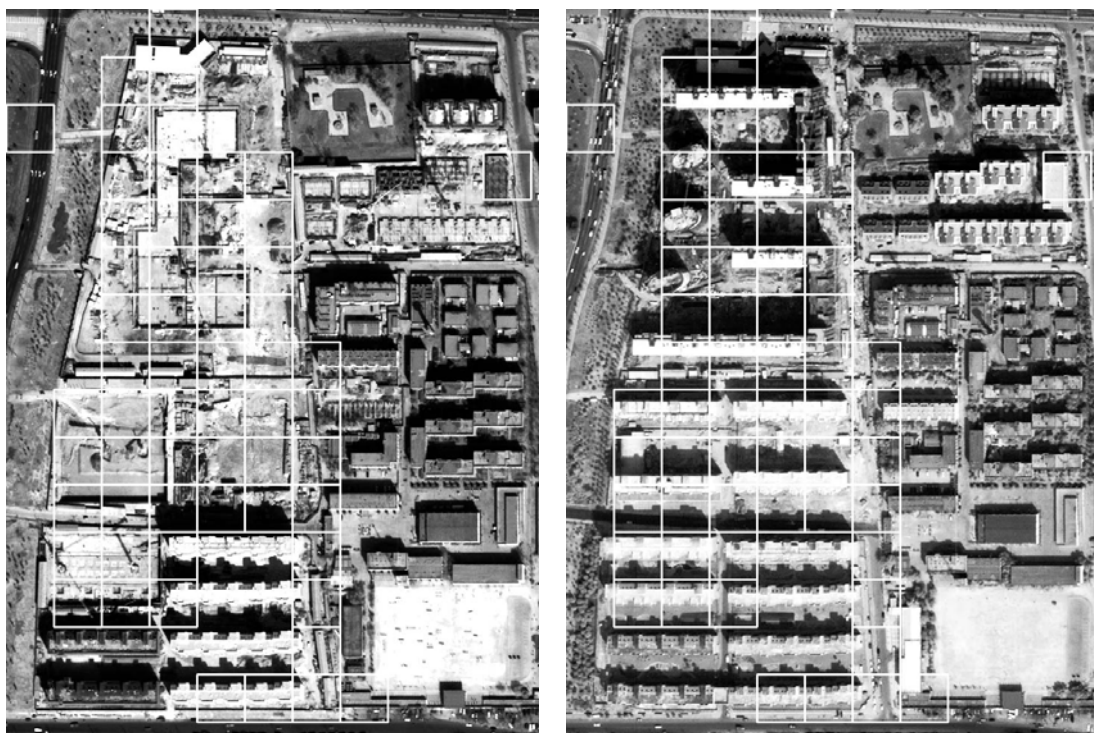


图 5-5 真实图像试验。图像大小：700x1000。左图为 2002 年图像；右图为 2003 年图像。格子大小：64x64。

5.8 小结

在这一章中，我们介绍了应用于结构物体变化检测的 DRF 模型，并且展示了初步的实验结果。变化检测 DRF 模型的核心思想是考虑图像在时间上的不连续性和在空间上的连续性，这两种性质分别是结构物体的变化和物体具有连续性事实的具体体现，它们是通过定义模型中的联合势和交互势而实现的。通过实验结果，我们可以看出变化检测 DRF 模型对简单图像的结构变化具有很好的检测能力，但是对于复杂的真实图像，并不能获得太好的结果，暂时还不能应用于真实图像的变化检测中。综合模型的各方面特性，变化检测 DRF 模型的局限性主要表现在：

✓ 元素选择的困惑

实验中选择元素是规则格子，选择这种元素的好处是处理起来简单，很容易定义元素之间的关系，并且由于格子具有一定的尺寸，因此对于图像间匹配的要求不高，特别是对于图像间的平移具有一定的容忍性。但是，我们最终要面对的是真实世界中的物体——比如建筑物，真实物体的形态很难简单用规则格子的组合来表示。规则格子的这种局限性在根本上决定了模型很难描述复杂的物体。另外一方面，使用规则格子作为基本元素需要我们预定义格子的尺寸，选择太大的格子可能导致格子内的特征不明显；太小的格子可能导致受噪声的干扰太大，同时也会增加可观的计算代价。就建筑物来说，由于在真实世界中呈现的尺度范围比较大，这就给格子尺寸的选择带来了困难；而且，若图像的分辨率发生了变化，也必须重新定义格子尺寸。

✓ 特征选择的局限性

基于规则格子进行特征提取，只能局限于格子内部的简单特征，很难提取出更高层次的特征——比如对建筑物来说非常重要的阴影特征。算法中计算的结构特征主要考虑的是直线相关的特征，虽然这些特征是建筑物的重要特征，但是它们只是必要条件，也就是说，通过这些特征我们无法区分其它具有这些特征的人造物体——比如道路，广场等。因此，我们变化检测的结果只能给出所有具有结构特征变化的物体。

✓ 参数优化的困难

从模型的定义中可以看出模型的复杂程度，参数优化需要计算模型的一阶和二阶导数，这需要预先使用解析数学的方法计算导数函数的形式，这是一个费力的过程。另外，我们的优化算法需要计算 Hessian 矩阵及其逆矩阵，这是一个费时的过程，而且优化过程中 Hessian 矩阵不一定一直具有可逆性，因此可能会导致优化算法的失败。

✓ 结果推断的困难

在结果推断过程中，面对的是一个组合优化问题，在算法中我们使用的是 ICM 算法。ICM 虽然具有很高的效率，但是即使具有很好的初始参数，依然无法保证能够找到全局极值，这就无法在理论上保证 ICM 推断出的标记配置是最优配置。另外，若图像的尺寸较大，相应的格子标记构成的组合空间尺度会显著增长，因此影响 ICM 推断出的最优配置结果。

✓ 训练集的生成问题

训练集的生成有两方面的问题：一是我们现在的数据库十分有限，只能提供 2002 和 2003 年的数据，无法生成一个较大的训练集，从而影响参数训练的结果；另一个问题是真实图像的训练集目前是通过手工标记的方法生成的，这是一个费力的过程。目前只能提供 2002 年的 GIS 数据作为先验知识，而无法提供 2003 年的 GIS 数据，因此也就无法利用 GIS 自动标记出那些变化元素。

尽管变化检测 DRF 模型具有一定的局限性，但是我们也应该看到它本身具有的能力和发展潜力。在今后的工作中，我们可以就其局限性研究相应的对策，对模型加以改进。具体可以考虑如下几种方向：尝试其它形式的元素——比如直线，区域等；依据物体的结构特性，引入更复杂的各向异性的邻域系统；寻求更好的参数优化方法，特别是对 Hessian 矩阵的估计方法；对于结果推断，可以根据模型特点选择尝试一些全局优化算法，比如 SA，GA 等；寻找更丰富的数据源尝试结合 GIS 自动生成训练集。

第六章 结束语

本章我们将对本文的工作及主要贡献进行总结，并对下一步的工作进行展望。

6.1 本文主要工作

本文在对以前遥感图像中物体提取与变化检测技术的分析和总结的基础上，结合项目的应用背景，以高分辨率 Quickbird 卫星图像数据为实验数据，围绕建筑物识别与变化检测这一主题进行了研究。本文主要工作及贡献如下：

➤ 提出了基于概率的建筑物轮廓识别算法

本文提出了一种建立在轮廓提取基础上，并结合概率模型进行建筑物轮廓识别的算法。在这种算法中，我们提出了剪切-融合算法用于轮廓的提取与合并，并建立了一个概率模型用以衡量某个轮廓区域是否为建筑物的可能性。这种算法的优点是能够通过概率模型把建筑物的各个特征结合起来，并通过学习由先验数据来决定模型参数，同时，模型具有很好的扩展性，新的特征很容易加入进来。相对于其它建筑物提取算法，本算法具有更好的鲁棒性和可扩展性。

➤ 提出了基于轮廓提取的建筑物变化检测算法

基于我们提出的建筑物轮廓识别算法，我们又提出了相应的建筑物变化检测算法。此算法的主要思想是首先找出不同时相图像中的建筑物轮廓，然后利用我们提出的概率模型比较对应的建筑物轮廓并给出它们变化的可能性，最后根据建筑物相应的变化概率值，判断它们是否发生了变化。通过实例可以看到这种算法对于一般纹理复杂度的建筑物的变化具有较好的检测能力，并且算法本身对图像的配准精度要求不高，具有较强的鲁棒性。

➤ 提出了变化检测 DRF 模型

我们提出了一种能够直接检测图像中结构物体变化的算法。基于前人提出的 DRF 模型，我们结合具体的应用背景，提出了用于结构物体变化检测的 DRF 模型。变化检测 DRF 模型的核心思想是考虑图像在时间上的不连续性和在空间

上的连续性，这两种性质分别是物体的变化和物体具有连续性事实的具体体现，它们是通过定义模型中的联合势和交互势而实现的。运用变化检测 DRF 模型，我们在不同图像上进行试验并获得了初步的实验结果，充分展示了这种算法的潜力和优势，为今后的工作奠定了基础。

6.2 未来工作展望

本文提出的一种建筑物轮廓识别方法和两种变化检测方法，都展示出了很好的性能或者巨大的潜力，也为今后的工作打下了良好的基础。未来的工作可以从如下方面展开：

➤ 基于概率的建筑物轮廓识别算法的改进

由于算法对现代化的更复杂的建筑物的识别效果不好，对图像的光照条件也具有一定的要求，而且目前选用的多边形近似算法并不十分理想。今后的工作中，我们可以通过使用更具有分辨力的特征、寻找更合理的特征空间和更好的融合算法、尝试更大的训练集并提高使用 GIS 进行参数训练的效果、寻找更好的多边形近似方法等途径对目前的算法进行改进。

➤ 基于轮廓提取的建筑物变化检测算法的改进

由于这种算法是基于建筑物轮廓识别算法的，因此建筑物轮廓识别算法在很大程度上决定了变化检测算法的效果。在以后的工作中，我们可以通过提高建筑物识别算法的正确率提高变化检测算法的正确率，另外，还可以通过比较两个不同时相图像的直方图消除它们之间的差别，使得应用旧图像上训练得到的优化参数在新图像上进行建筑物识别能获得更高的正确率。

➤ 变化检测 DRF 模型的改进

通过第五章的分析可知，我们提出的变化检测 DRF 模型还具有很大局限性，目前还无法在真实图像试验中取得很好的实验结果。但是，我们也应该看到这个模型本身蕴涵的巨大潜力，在今后的工作中，我们可以就其局限性研究相应的对策，对模型加以改进。具体可以考虑如下几个方向：尝试其它形式的元素；引入更好的邻域系统；寻求更好的参数优化和结果推断方法；训练集的自动化生成等。

参考文献

- [1] H.Mayer, “Automatic Object Extraction from Aerial Imagery - A Survey Focusing on Buildings”, Computer Vision and Image Understanding, vol.74, 1999, pp. 138-149.
- [2] Wenbo Song; Haithcoat, T.L., “Development of comprehensive accuracy assessment indexes for building footprint extraction”, Geoscience and Remote Sensing, 2005, vol.43, Issue 2, pp.402 - 404.
- [3] A.Baumgartner, C.Steger, H.Mayer, “Multi-resolution, semantic objects, and context for road extraction”, in Proceedings, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, pp. 140-156, 1997.
- [4] M.de Gunst and G.Vosselman, “A semantic road model for aerial image interpretation”, in Proceedings, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, pp. 107-122, 1997.
- [5] J.A.Shufelt, “Exploiting photogrammetric methods for building extraction in aerial images”, in Proceedings, International Archives of photogrammetry and Remote Sensing, 1996, Vol. (31) B6/VI, pp. 74-79.
- [6] X.Wang, J.Lim, R.T.Collins, and A.R.Hanson, “Extracting surface textures and microstructures from multiple aerial images”, in Proceedings, Computer Vision and Pattern Recognition, 1997, pp. 71-76.
- [7] O.Henricsson, “The role of color attributes and similarity grouping in 3-D building reconstruction”, Computer Vision and Image Understanding, 72(2), 1998, 163-184.
- [8] A.Fischer, T.H.Kolbe, “Extraction buildings from aerial images using hierarchical aggregation in 2D and 3D”, Computer Vision and Image Understanding 72(2), 1998, 185-203.
- [9]K.Kulschewski, “Building recognition with Bayesian networks”, in Proceedings, Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, 1997, pp. 196-210.
- [10] R.Ruskone, L.Guigues, “Vehicle detection on aerial images: A structural approach”, in Proceedings, 13th International Conference on Pattern Recognition, 1996, Vol. III, pp. 900-903.

- [11] R.Mohan and R.Nevatia, “Using perceptual organization to extract 3-D structures”, IEEE trans. Pattern Analysis and Machine Intelligence. 11(11), 1989, 1121-1139.
- [12] M.Herman and T.Kanade, “The 3D MOSAIC scene understanding system: Incremental reconstruction of 3D scenes from complex images”, in Proceedings, Image Understanding, 1984, pp. 137-148.
- [13] U.Weidner, “Digital surface model for building extraction”, in Proceedings, Automatic Extraction of Man-made Objects from Aerial and Space Images(II), pp. 193-202, 1997.
- [14] Krishnamachari, S., Chellappa, R., “Delineating buildings by grouping lines with MRFs”, IEEE Trans. Image Processing, vol. 5, Issue 1, Jan. 1996 pp.164 - 168.
- [15] Gerke, M., Heipke, C., Straub, B.-M., “Building extraction from aerial imagery using a generic scene model and invariant geometric moments”, in Proceedings, Remote Sensing and Data Fusion over Urban Areas, 2001, pp.85 - 89.
- [16] Garcin, L., Descombes, X., Le Men, H., Zerubia, J., “Building detection by Markov object processes”, in Proceedings, International Conference on Image Processing, 2001, Volume 2, pp.565 - 568 vol.2.
- [17] A.Stassopoulou, T.Caelli, “Building Detection using Bayesian Networks”, International Journal of Pattern Recognition and Artificial Intelligence, 14(6):715-733.
- [18] Lin C.; Nevatia R., “Building Detection and Description from a Single Intensity Image”, Computer Vision and Image Understanding, 1998, vol.72, no. 2, pp. 101-121(21).
- [19] Haithcoat, T.L.; Song, W.; Hipple, J.D., “Building footprint extraction and 3-D reconstruction from LIDAR data”, in Proceedings, Remote Sensing and Data Fusion over Urban Areas, 2001, pp.74 - 78.
- [20] Yanfeng Wei; Zhongming Zhao; Jianghong Song, “Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection”, in Proceedings, Geoscience and Remote Sensing Symposium, 2004, vol.3, pp.2008 - 2010.

- [21] Tison, C.; Tupin, F.; Maitre, H., "Retrieval of building shapes from shadows in high resolution SAR interferometric images", in Proceedings, Geoscience and Remote Sensing Symposium, 2004. vol.3, pp.1788 - 1791.
- [22] Yanlin Guo, Sawhney, H.S., Kumar, R.; Hsu, S., "Learning-based building outline detection from multiple aerial image", in Proceedings, Computer Vision and Pattern Recognition, 2001, vol.2, pp.II-545-II-552.
- [23] 彭文辉, 基于信息熵的合成孔径雷达图像水体变化检测, 硕士学位论文.
- [24] Ashbindu Singh, "Digital change detection techniques using remotely-sensed data," International Journal of Remote Sensing, vol.10, NO.6, 989-1003, 1989.
- [25] Tlidd W J, "Urban and regional land use change detected by using landsat data", Journal of Research U.S.G.S, 5(5): 529-534, 1997.
- [26] Mark J.Carlotto, "Detection and analysis of change in remotely sensed imagery with application to wide area surveillance," IEEE Trans. Geosci.Remote Sensing, vol.6, NO.1, 189-202, 1997.
- [27] 章孝灿等, 遥感数字图像处理, 1997, 浙江大学出版社.
- [28] Jianwen Ma, Li Qiqing, "The endangered rare plant coverage change detection in twelve years by using TM/ETM data," in Proceedings, International Geoscience and Remote Sensing Symposium, 2003, vol.3, pp.2008 - 2010.
- [29] Nielsen A A, Conradsem K, Simpson J J, "Multivariate Alteration Detection (MAD) and MAF postprocessing in multitemporal image data: new approaches to change detection studies," Remote Sensing Environment, 64:1-19,1998.
- [30] R.D.Johnson and E.S.Kasisichke, "Change vector analysis: a technique for the multispectral monitoring of land cover and condition," INT.J.Remote.Sens. vol.19, no.3, 411-426, 1998.
- [31] 廖明生, 朱攀, 龚健雅, "基于典型相关分析的多元变化检测," 遥感学报, 第4卷第3期, 197-201, 2000.
- [32] Peng WH, Pan CH, V.Prinet, "Unsupervised change detection for flood in SAR images", in Proceedings, Asian Conference on Computer Vision, 2004, Korea.
- [33] L.Bruzzzone, "Automatic analysis of the difference image for unsupervised change detection," IEEE Trans. Geosci. Remote Sensing, 2000, vol.38, no.3, 1171-1182.

- [34] L. Bruzzone and D. Fernandez, "An adaptive parcel-based technique for unsupervised change detection", *Int. J. Remote. Sens.* 2001, vol. 21, no. 4, pp. 817-822.
- [35] T. Yamamoto and H. Hanaizumi, "A change detection method for remotely sensed multispectral and multitemporal images using 3-D segmentation", *IEEE Trans. Geosci. Remote Sensing*, 2001, vol. 39, no. 5, pp. 976-985.
- [36] L. M. T. Carvalho, L. M. G. Fonseca, "Digital change detection with the aid of multiresolution wavelet analysis," *Int. J. Remote. Sens.* 2001, vol. 22, no. 18, pp. 3871-3876.
- [37] T. Kasetkasem and P. K. Varshney, "An image change detection algorithm based on markov random field models," *IEEE Trans. Geosci. Remote Sensing*, vol. 40. no. 8, 1815-1823, 2002.
- [38] Jordi Inglada, "Change detection on SAR images by using a Parametric Estimation of the Kullback-Leibler Divergence," in *Proceedings, International Geoscience and Remote Sensing Symposium*, 2003, vol. 6, pp. 21-25.
- [39] Y. Delignon, R. Garello, A. Hillion, "Statistical modelling of ocean SAR images", *IEE proc. on Radar, Sonar and Naving*, 1997, vol. 44, no. 66, 348-354.
- [40] Watanabe, S.; Miyajima, K.; Mukawa, N., "Detecting changes of buildings from aerial images using shadow and shading model", 1998, in *Proceedings, 14th International Conference on Pattern Recognition*, pp. 1408 - 1412, vol. 2.
- [41] Huertas, A.; Nevatla, R., "Detecting changes in aerial views of man-made structures", 1998, 6th *International Conference on Computer Vision* pp. 73 - 80.
- [42] D. Gerson and S. Wood, *RADIUS Phase II: "The RADIUS Testbed System"*, Image Understanding, Monterey, CA, 1994.
- [43] Agouris, P., Mountrakis, G. and Stefanidis, "Automated Spatiotemporal Change Detection in Digital Aerial Imagery", *SPIE Proceedings Vol. 4054: Aerosense*, 2000, pp. 2-12, Orlando.
- [44] Shintaro, Miyajima, K., "Detecting building changes using epipolar constraint from aerial images taken at different positions", 2001, in *Proceedings, International Conference on Image Processing*, pp. 201 - 204 vol. 2.

- [45] Tuong Thuy Vu; Matsuoka, M.; Yamazaki, F., “LIDAR-based change detection of buildings in dense urban areas”, in Proceedings, Geoscience and Remote Sensing Symposium, 2004, pp.3413-3416 vol.5.
- [46] Sanjiv Kumar, Martial Hebert, “Discriminative Random Fields: A Discriminative Framework for Contextual Interaction in Classification”, in Proceedings, International Conference on Image Processing, 2003.
- [47] 边肇祺, 张学工等, 模式识别, 清华大学出版社, 1999
- [48] Thomas P. Minka, “Algorithms for maximum-likelihood logistic regression”. Technical Report. Stats 758, Carnegie Mellon University, 2001.
- [49] J. Brian Burns, Allen R. Hanson, Edward M. Riseman, “Extracting straight lines”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.8, no.4, p.425-455, 1986.
- [50] 郎锐, 数字图像处理学, 北京希望电子出版社, 2002.
- [51] 袁卫等, 统计学(修订本), 中国统计出版社出版, 1994.
- [52] M. Visvalingam and J. D. Whyatt, “The Douglas-Peucker Algorithm for Line Simplification: Re-evaluation through Visualization”, in Proceedings, Computer Graphics Forum, vol.9, Issue 3 (1990), pp.213-228.
- [53] S.Z.Li, Markov Random Field Modeling in Computer Vision, Springer, 1995.
- [54] D. M. Greig, B. T. Porteous, and A. H. Seheult. “Exact maximum a posteriori estimation for binary images”. Journal of Royal Statis. Soc., 51(2):271-279, 1989.
- [55] C. Fox and G. Nicholls. “Exact map states and expectations from perfect sampling: Greig, porteous and seheult revisited”, in Proceedings, 20th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Sci. and Eng., 2000.
- [56] J. Besag. “On the statistical analysis of dirty pictures”, Journal of Royal Statistical Soc., B-48:259-302, 1986.
- [57] S. Geman and D. Geman, Stochastic relaxation, “Gibbs Distribution and Bayesian Restoration of Images”, IEEE Trans. Pattern Recognition and Machine Intelligence, vol.6, no.6, Nov. 1984.

致 谢

我在攻读硕士期间从事的学习和科研工作，是在导师和实验室各位老师、同学的指导和帮助下完成的。在完成论文之际，我谨表最诚挚的谢意。

首先感谢我的导师普林特副研究员。在研究生三年的学习和工作中，她给予了我精心的指导和尽心的关怀。普林特不仅学术精深、工作严谨，而且平易近人、为人师表。她对我的影响将使我终生受益。

感谢胡占义研究员，王润生教授，潘春洪副研究员，杨青研究员对我论文的指导与帮助。

感谢我们组的吴刚，段静辉，任坤，张艳洁，余战武，洪碧珍，白俊，白笛暮，张朝晖，彭婷。感谢中法实验室的腾军，杨余久，陈其兵，刑宏杰，尹巍巍，陈翀，曲寒冰，程章林，王勇。

感谢模式识别国家重点实验室给我提供了良好的学习环境和实验条件。感谢娄文利老师、连国臻老师，研究生部的李磊、卜树云等老师，为我们的生活提供的帮助。

最后我要感谢我的爸爸、妈妈和女朋友张仲，感谢他们对我至始至终的支持和鼓励！感谢我的那些朋友们！

刘炜

2005 年 5 月

作者个人简历

刘炜，男，汉族，团员，1980年8月21日出生于湖北省武汉市。1998年7月毕业于北京一零一中学，同年9月考入北京大学电子学系学习，并于2002年7月毕业，获理学学士学位。2002年9月至2003年6月，在中国科学院研究生院学习研究生基础课。2003年8月至2005年6月在中科院自动化所模式识别实验室攻读硕士学位，主要研究方向为高分辨率卫星图像中建筑物的识别和变化检测。

硕士期间提交及发表的论文：

[1] Liu Wei, Veronique Prinet, “Building Detection from High-resolution Satellite Image Using Probability Modeling”, International Geosciences and Remote Sensing Symposium (IGARSS2005), Seoul, Korea. (accepted)