

学校编码: 10384

分类号_____密级_____

学 号: 200228044

UDC_____

厦 门 大 学
硕 士 学 位 论 文

建筑目标识别中
检测与分割技术的研究

Building Detection and Extraction in
Recognition System

叶 聪 颖

指导教师姓名: 李翠华 教授

专 业 名 称: 计算机应用技术

论文提交日期: 2005 年 5 月

论文答辩日期: 2005 年 6 月

学位授予日期: 2005 年 月

答辩委员会主席: _____

评 阅 人: _____

2005 年 5 月

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文而产生的权利和责任。

声明人（签名）：

年 月 日

摘要

计算机视觉技术的迅速发展使得自动目标识别在当今的工业生产、日常生活娱乐以及国防军事等领域中的作用日益显著，应用前景越来越广泛。作为自动目标识别的组成部分之一，建筑物目标的提取与识别对于城市测绘与统计规划、三维数字地图重建、导航控制等诸多方面有着重要的意义。但是目前建筑目标提取与识别的研究工作中还存在许多难以解决的问题，是一个具有挑战性的课题。

在建筑目标识别系统中，迅速地检测以及分割提取出图像中的主要建筑目标候选区域非常重要，它能缩小搜索范围并且有助于提高识别精度。但是由于受到天气、光线、建筑物密度等众多因素的影响，分割提取工作也具有相当的难度。本文针对基于单幅图像的建筑区域检测以及分割提取问题进行研究。

在建筑物区域检测中，本文引入多分辨率分析的思想，结合小波变换的提升格式形成一种分级处理的方法，既能排除局部噪声影响，又能大大加快处理速度。在图像预处理阶段，我们考察了现有的对比度拉伸展宽方法然后通过改进方法获得了更好的效果，并且结合 Sobel 算子和一系列细化、断线修补和小线段剔除的处理流程从图像中较好地提取边缘。本文假定输入图像都是基本正立的，在此基础上提出了一种基于竖直线段的空间位置关系模式实现分组并形成建筑物粗分割和目标候选区域提取的方法。

本文还结合了视觉注意力机制开展针对人工建筑目标的主动视觉搜索方面的研究，根据应用特点对已有的显著性模型中的特征提取和规格化方法进行改进，使之更适合于建筑物区域检测的任务。

关键词：建筑目标识别；机器视觉；视觉显著性

Abstract

With the rapid development of computer vision technology nowadays, automatic target recognition is playing a more important role in industry, entertainment, military realm and many other fields. Being a part of automatic target recognition technology, extraction and recognition technique for building target is significant for applications such as city statistics, reconstruction of 3D maps and steering control system. However, there are still many unresolved problems in extraction and recognition of buildings, which make researches in this area challenging and rewarding.

Effective detection and extraction of possible building areas in the images is very important for the building recognition system. It helps to reduce the searching area and improve the precision of recognition. However, the influence of weather, illumination, building density and many other factors have made extraction a quite difficult task. In this paper we introduce the research on the detection and extraction of buildings from monocular images.

In building detection, our approach uses multi-resolution analysis and lifting scheme of wavelet transform to form a processing architecture. It is capable of removing noise effect and speeding the whole process. In the step of image preprocessing, we first analyze existing contrast extending methods and then describe our method of improvement. After this step, we extract the desired edges more effectively by using Sobel edge detector as well as a serial of processes, containing edge thinning, gap filling and trivial edge removal. Moreover, on the hypothesis that images are upright, we propose a building detection and extraction method based on straight edge grouping technique.

We also apply the visual attention mechanism to the problem of building detection. According to the detail studies of building images, we improve the methods for feature extraction and normalization in current saliency model,

making them more appropriate for building detection task.

Keywords: Building Recognition; Computer Vision; Visual Saliency.

厦门大学博硕士学位论文摘要库

第一章 绪论	1
1.1 引言	1
1.2 建筑目标识别的研究背景与研究内容	1
1.3 本文主要研究内容	3
第二章 建筑目标识别的研究现状与相关技术	5
2.1 引言	5
2.2 国内研究现状	5
2.3 国外研究现状	6
2.4 主要技术难点	9
2.5 建筑目标识别系统的构成	10
2.6 小结	12
第三章 基于竖直线特征的建筑物区域检测与分割	13
3.1 引言	13
3.2 多分辨率分析	13
3.2.1 图像金字塔	14
3.2.2 小波变换	15
3.2.3 小波变换的提升格式	18
3.2.4 多分辨率分析在建筑物区域检测中的应用	20
3.3 图像预处理	22
3.3.1 对比度展宽	22
3.3.2 实验结果	24
3.4 边缘提取和竖直线提取	25
3.4.1 基于 Sobel 算子的边缘提取	25
3.4.2 竖直线提取	28
3.4.3 实验结果	28
3.5 边缘后处理	30
3.5.1 断线修补	31
3.5.2 小线段剔除	33
3.5.3 实验结果	34
3.6 竖直线分组与建筑物分割	36
3.6.1 分组时空间位置关系的计算	37
3.6.2 分组算法	40
3.6.3 分组合并与筛选	41
3.6.4 实验结果	43

3.7 小结	46
第四章 基于视觉注意力机制的建筑物区域检测	49
4.1 引言	49
4.2 视觉显著性模型	49
4.3 显著性建筑物区域检测	52
4.3.1 特征提取	52
4.3.2 显著图合成	54
4.3.3 实验结果	56
4.4 小结	58
第五章 结束语	59
5.1 主要研究工作总结	59
5.2 展望	61
参考文献	63
研究生期间的研究成果	67
致 谢	69

第一章 绪论

1.1 引言

伴随着计算机硬件技术和一些相关软件技术的不断升级，计算机视觉技术也获得了空前地发展，应用领域已经遍及遥感图像处理、医学图像处理和理解，目标识别、控制和导航、自动视频监控跟踪、人机交互界面、基于内容的图像视频检索以及虚拟现实等众多方面。自动目标识别(ATR: Automatic Target Recognition)是计算机视觉研究中一个重要的组成部分，在工业生产、国防军事、日常生活和娱乐中都有着越来越广泛的应用前景，本文针对自动目标识别领域内的建筑目标识别这一课题开展研究。该研究围绕着 ATR 实验室第六研究室的图像与视频信息处理任务展开，并且得到了国家自然科学基金和福建省自然科学基金的支持。

1.2 建筑目标识别的研究背景与研究内容

建筑目标检测与识别是 ATR 技术的组成部分之一。ATR 技术的研究工作开始于 20 世纪的中前期，至今仍是一个研究热点和攻关性研究课题。ATR 技术应用很广泛，包括军事上的武器制导、电子侦察，民用领域中的安全监控、自动控制、工业零件检测等等，研究内容也涵盖了模型检测匹配、机器学习、多传感器信息融合、嵌入式芯片技术等诸多方面。目前 ATR 领域的一个重要特点就是问题的针对性强，尚不可能存在通用的识别方法，只能面对具体任务来寻找专用技术和方法。

作为 ATR 组成部分之一的建筑目标检测与识别研究开始于 20 世纪 80 年代初，是近 30 年来一个活跃的研究领域，它不仅在军事上有着重要的意义，在民用经济中也发挥着越来越大的作用，随着数字化城市进程的不断推

进，应用越来越广。对城市测绘来说，航拍图像中建筑目标的自动或半自动探测和提取能极大地减少人工的工作量，提高工作效率，达到获悉城市建筑群分布，为城市统计和规划提供支持手段，提取的信息还可用于地面建筑物分布图的自动测绘，并结合计算机图形技术和虚拟现实技术实现三维数字地图重建。另外建筑目标识别技术与车载或手持移动设备相结合还可以在城市交通、旅游导航等领域中开辟一番新的天地。总之，建筑目标识别的研究对于国防建设和民用科技都很有意义，应用前景越来越广泛。

建筑目标的检测与识别涉及到计算机视觉、模式识别、图像处理以及人工智能等领域。要实现建筑目标识别需要结合多种信息，对硬件软件也有较高的要求。由于对建筑物的拍摄受到天气、时间等因素的限制，仅从光学图像很难达到准确搜索的目的，而且光学图像的分辨率对建筑目标检测识别也有影响。在高分辨率的图像中每个像素代表的距离小，在低分辨率的图像中每个像素代表的距离大，因此图像分辨率对准确定位有一定的影响。另外拍摄视角对建筑目标识别也有影响，仅从二维图像信息识别建筑目标缺少景深数据，存在着信息量不足的缺点，因此准确搜索和识别建筑目标需要多源信息融合技术。在图像中对建筑物进行分割也是一个很难处理的问题，尤其是在城市建筑物图像中，建筑物之间的距离比较近并且外形、颜色各方面多有雷同，分割定位以及识别的难度都很大。

目前在建筑目标识别方面最主要的研究内容是航拍、遥感和SAR等图像中建筑物的半自动或自动检测、提取和识别^[1]，其中又以检测和提取方面的研究居多，识别方面的研究难度比较大，并且需要以检测和提取为基础，因此目前进行得相对较少。文献[2]中介绍的是一种基于侧面轮廓信息的建筑物识别在手持自动旅游导航系统中的应用，其它未见到相关的研究报道。这方面研究的难度比较大，研究成果可以结合俯视的航拍图像实现全方位的建筑提取和识别，辅助实现城市建筑三维数字化的工作，完成GIS建库等，可以极大地提高工作效率，也可以用于低空飞行器的辅助导航控制或者手持的自

动旅游导航等应用领域。

1.3 本文主要研究内容

本文的主要研究对象是中远距离拍摄的建筑物可见光图像，这一点和目前通常研究的接近正俯视的航拍图像有所区别。研究的目的是探索一种基于建筑物侧面轮廓信息进行建筑区域检测、分割和目标粗定位，并利用事先建立的目标模板实现有效的建筑目标识别的方法。要达到准确识别建筑目标的要求，首先需要快速地提取图像中的建筑物区域以及分割和定位出主要的建筑物所在方位，本文阐述的就是整个建筑目标识别系统中建筑区域检测以及主要建筑目标候选区域粗分割和定位方面的研究工作，它是实现建筑目标识别任务的重要基础，能大大地缩小搜索范围，提高搜索效率和准确性。

本文共分五章，具体组织如下：

第一章是绪论，简要介绍了建筑目标识别研究的背景、主要内容和应用领域。

第二章是建筑目标识别的研究现状与相关技术，主要介绍建筑目标识别领域当前研究的各个方向，国内外在这些方面的研究进展情况，采用的相关技术等，并且分析了我们要研究的课题所面临的主要技术难点以及相应的系统构成。

第三章是基于竖直线特征的建筑物区域检测与分割的研究，主要介绍了本文基于建筑物竖直线特征进行建筑物区域检测的工作以及在区域检测的基础上如何对主要建筑物进行分割和粗定位的研究，并给出了相关的实验结果。

第四章是关于结合视觉注意力机制进行建筑物区域检测的研究，介绍了视觉注意力机制等主动视觉方面的当前研究状况，阐述了本文结合建筑物区域应用所做的改进工作以及实验结果分析。

第五章是结束语，总结了本文的主要工作和不足之处。

第二章 建筑目标识别的研究现状与相关技术

2.1 引言

目前建筑目标识别领域的研究主要集中在航拍图像中建筑物的提取、识别以及相应的三维建模,研究成果应用于地图特征提取、城市建筑自动测绘、城市人口调查、地理信息系统建库等领域,对军事、民用都具有重要意义,相关的研究论文也频频发表在各种高水平的学术会议和刊物上。总的来说,建筑物的提取还是一个尚未完全解决的学术难题,很多系统都是建立在半自动检测和少量人工干预的基础上,完全自动的检测也是针对建筑物比较稀疏的地区才比较有效。提取建筑物所采用的方法多数是基于建筑物的二维和三维特征,也有借助颜色信息或结合DEM数据等。这方面的研究方法大致可分为基于边缘检测、区域分割以及立体模型三类。典型的研究内容有基于线条提取分析^[3]和知觉分组技术^[4]并利用阴影和透视几何规则等辅助信息^[5]进行建筑物检测和提取,结合DEM数据的三维建筑物建模^[6],构建基于知识的提取系统^[7],通过影像测量并结合物体的几何知识构建出多面体对象模型^[8],结合已有的二维地图矢量数据利用航空激光扫描^[9,10]或激光高度计数据^[11,12]进行建筑物建模,融合三维深度传感信息、多CCD相机和彩色高分辨率数码相机获取的数据实现建筑物建模^[13]等。

以下针对国内和国外的研究情况分别做比较详细地介绍。

2.2 国内研究现状

国内在建筑物提取和识别方面的工作开展得比较晚,基本上是在国外相关领域的研究成果基础上发展起来的,但近几年的进展比较快。不少高校都针对航拍图像中建筑物的提取和三维模型重建开展了一些研究工作,提出了

不少改进方法和新思路，既有在半自动提取方面的研究^[14,15]，也有针对全自动提取的研究^[16,17]，研究的方法主要有基于直线提取分析以及灰度信息的，也有基于模板^[18]或其它一些方法^[15]的建筑物提取。

西北工业大学的杨益军和赵荣椿教授在文献[16]中讨论了一种在航空图像中自动检测建筑物目标的方法，他们先对建筑物目标进行了分割，并利用建筑物的形状特征和阴影信息进行了验证。在文献[17]中作者提出了一种新型的处理方法，在边缘提取的基础上经过轮廓跟踪并采用 *splitting* 方法来提取直线，之后引入一系列直线处理方法来弥补直接提取的不足，文中还借助几何结构元分析的方法最终形成图像中建筑结构的提取。

另外国内也在三维建模方面开展了一些研究工作^[19]，但是在基于侧面轮廓的建筑物提取以及建筑物识别方面尚未见到相关的研究成果发表，属于一个空白的研究领域。

2.3 国外研究现状

国外尤其是美国对建筑物检测提取和识别领域的研究从二十世纪八十年代就开始了，目前在这方面的研究也处于领先地位。航拍图像中建筑物的提取最初采用的是区域生长结合简单的图像几何模型的方法，后来的工作逐步使用了阴影分析以及线条、角点分析来推导建筑物的结构，近年来利用基于数据的多模式分析来推导建筑物 2D 和 3D 模型的方法渐渐流行起来。美国卡内基梅隆大学的 Mapslab 小组，南加州大学的计算机视觉小组，马萨诸塞州大学的计算机视觉实验室，德国慕尼黑 University of the Federal Armed Force 的 Mayer 领导的小组等实验室都在航拍图像建筑物提取研究领域做了大量的工作。

卡内基梅隆的 Mapslab 小组相继完成了四个建筑物提取的系统: BUILD, BUILD+SHAVE, VHBUILD 和 PIVOT, 以下分别做简要介绍。

1. BUILD 系统假设所得到的图像是接近正俯视角度的，即所有建筑是

近似盒子形状的。该系统基于线条和角点的信息来分析建筑物。并通过寻找边缘的高曲率点作为拐角点，把分段的边缘连接起来构成盒子框架。该系统没有产生建筑物的三维模型。

2. **BUILD+SHAVE** 系统是一个杂交系统，它包含了 **BUILD** 系统以及基于阴影的校正系统 **SHAVE**。该系统基于与 **BUILD** 系统一样的假设，它使用了一般的阴影阈值来描述阴影区域，并使用已知的光线入射角，结合阴影信息计算建筑物高度。该系统能输出产生三维建筑物框架。
3. **VHBUILD** 系统在前者的基础上降低了对获取图像的限制，但结合了成像模型。该系统能提取具有水平屋顶和人字形尖屋顶等特征的房屋，并对屋顶进行建模。系统对建筑物的线与线之间的角度没有限制，能处理一定程度的几何变换，并且使用了消失点来修正建筑物“盒子”和角点。该系统还使用了与 **SHAVE** 相同的阴影检测法，最终能产生三维的建筑模型。
4. **PIVOT** 使用摄影测量方法进行特征提取，该系统使用了一种改进的消失点探测方法来检测水平，垂直，倾斜的消失点，然后用这些信息来探测建筑物的结构。该系统同样能产生三维的建筑模型。

南加州大学最初的工作^[20-23]是从平顶房屋的提取开始的，并且假设建筑物的形状都是一些长方体的组合（没有圆形和斜面等），但航拍的视角可以倾斜。在此假设的基础上，首先对图像进行线条检测，然后结合拍摄视角的几何变换和投影规则形成关于建筑物屋顶的假设。在这些初步形成的假设中有很多是虚假的目标，因此接下来就要结合提取出来的建筑物影子信息和侧面墙体信息对这些假设进行单独的验证筛选，淘汰前面很多不正确的假设。这个初步的验证工作是针对每个假设独立进行的，并没有考虑它们之间的相互关系，这样导致有一些不合理的假设仍然通过了这个初步的验证，所以系统中进一步使用了建筑物在三维空间上的相互包含和重叠关系所应该遵守

的几何规则来做进一步的验证，达到进一步排除不合理假设的目的。这种方法在处理各种航空图片时建筑物的检出率能达到 70% 以上，为了获得更好的效果，还可以采用多张不同角度的图像融合处理或者在处理过程中引入少量人工干预的方法。在以上这些成果的基础上，南加州大学又有了更深入的研究，他们利用多幅不同角度的航空照片来提取具有较复杂屋顶和形状的房屋以及融合多种信息进行处理^[24]。

马萨诸塞州大学的计算机视觉实验室也是和美国国防部等相关部门合作，对航拍图像中建筑物的识别和重建进行了深入地研究。其 ASCENDER2 和 RADIUS 两个项目取得了很大的成功，也已经投入实际应用。在 RADIUS 系统中，首先针对航拍图像中的低级特征进行提取和分组，辅助屋顶多边形的提取。除了这种自底向上的方法外，系统还结合自顶向下的方法来验证特征以及特征之间的关系，达到排除错误、虚假特征的目的。该系统还采用了一些图论的处理方法，通过对图中环路的处理来提取可能对应的屋顶多边形结构。为了得到更精确细致的建筑物三维模型，该系统还对航拍图像中的纹理进行了分析，克服了早期方法没有对可能包含大量重要信息的纹理结构进行研究的缺陷，通过对同一地区的多幅航拍图像应用复杂的纹理映射技术，进行纹理和微结构分析，来构造建筑物三维模型的纹理，从而为更准确的识别建筑物提供了保证。

近年来的研究趋势表明国外在航拍图像中建筑物提取、识别和重建方面的研究已经由原来的单一图像提取^[25,26]转向利用建筑物的多角度、多时段图像以及多传感器融合的提取^[27-30]，并且更多地利用先验知识以及考虑更加细化的建筑物特征来建立更复杂的模型。最新的研究进展还表明，利用地面拍摄的建筑物信息融合航拍图像中提取的建筑物信息可以用来构造更精确的三维模型。

在基于侧面轮廓的建筑目标提取和识别方面，国外的研究也相对较少，目前见到相关成果发表的只有德国斯图加特大学所做的一些工作^[2]，他们开

发了一种手持的旅游导航辅助系统，它首先使用GPS获得一个大致的定位，其次再针对拍摄的建筑物可见光图像与系统中事先存储的建筑物模型进行形状匹配，从而告诉使用者当前所见到的建筑物名称以及相关的各种信息。他们在进行形状匹配时采用的是广义Hough变换技术，而对于事先存储的建筑物轮廓模型则是采用手动CAD建模得到的，另外GPS的应用也使得搜索范围大大缩小，也就是说他们通过各种辅助手段，极大地降低了建筑目标识别的难度，简化了检测、提取和识别所需要处理的问题，因此系统的效果能达到一个相对较好的程度。

2.4 主要技术难点

一般针对航拍图像的建筑目标提取与识别研究工作所遇到的难点既有来自天气、成像等因素的干扰，也受到建筑密度、建筑结构复杂程度的影响，还有拍摄的倾斜角度以及高分辨率图像中过多的细节信息也带来很多困难。

本文研究的实验对象是中远距离拍摄的建筑物侧面可见光图像，它与通常基于航拍图像的研究有一些共同之处，更有很多各自不同的特点。首先由于距离较远且受到光照、雾气、噪声、成像比例等因素的影响，建筑目标的边缘一般较为模糊，难以完整提取。受光照的影响，目标建筑上还可能有阴影，从而造成虚假边缘。在城市建筑物密集的区域，目标建筑可能淹没在周围的建筑群中，众多细小杂乱的干扰边缘以及过多的局部细节也给处理工作带来很大的困难。

在边缘提取的基础上要进行建筑区域的分割定位处理。城市建筑目标的分割本身就是一个尚未完全解决的问题，在复杂背景下的图像分割尤其困难，目前不存在通用的解决方法，只能针对具体问题采用特定的规则和方法来解决。由于城市建筑的密集程度较高，外观上还比较相似，建筑物自身可能存在多种色彩以及不同建筑物也可能具有相似或雷同的颜色，因此在缺少深度信息和大量先验知识的情况下很难进行精确的分割，但是如果能够实现建

筑目标候选区域的粗分割，缩小进一步搜索的区域，对于识别任务也是非常有意义的。

我们的建筑目标识别系统在进行建筑识别工作时首先采用的是基于建筑物外轮廓的匹配，所以轮廓的提取对后继的工作也就非常重要。如果建筑物和背景能被成功分割并形成二值图像，那么外轮廓的提取就相对容易，但由于精确的分割就是一个难以解决好的问题，所以只能基于边缘结果来提取外轮廓，那么建筑物的窗、纹理等形成的边缘就很难被去除，同时建筑物外部墙体反光等原因也会导致轮廓边缘提取不完全。

2.5 建筑目标识别系统的构成

本文研究的内容是基于侧面轮廓信息的建筑物提取及识别，它与常见的航拍图像中建筑目标提取和识别任务有很多不同之处，因此系统的处理步骤和流程也各不相同。我们在针对研究内容进行分析之后设计了如下的系统结构，它包括两大部分，一部分是目标建筑物模板的建立，另一部分是根据模板在输入图像中进行目标搜索。

建立目标建筑物模板的部分包含以下四个阶段

1. 建筑区域检测

人工建筑物区别于其它自然景物的一个重要特征就是直线段，我们假设所处理的图像都是基本正立的，也就是说建筑物的外部墙体就对应着竖直方向的直线段，即墙是建筑物的重要证据，因此对建筑区域的检测可以归结为对竖直线段的检测，拥有大量竖直线段的区域可能就是建筑物密集的区域。

2. 建筑物分割和粗定位

在竖直线段提取的基础上，计算竖直线之间的空间位置关系，然后依照预先设定的一些模式和规则进行分组，最后再对各个分组进行验证与合并，通过这些处理得到的每一个分组都对应着一个建筑物的候选区域。

3. 手工指定目标建筑物的建模区域并在指定区域内进行外轮廓提取

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库