

青岛大学

硕士学位论文

基于手持设备的增强现实技术研究与应用

姓名：隋毅

申请学位级别：硕士

专业：计算机软件与理论

指导教师：邵峰晶

20090603

摘要

增强现实（Augmented Reality，简称 AR）技术将计算机生成的信息同真实世界的场景进行融合，在三维空间中实现二者位置的“对齐”，并且具有实时交互的特点。AR 在军事、医疗、教育、文化以及娱乐等领域具有广泛的应用前景。随着移动技术的发展和普及，具有体积小、携带方便、功耗低等特点的手持设备（智能手机、PDA 等）已经成为 AR 系统良好的实现平台。因此，研究基于手持设备的增强现实技术具有重要的理论意义和实用价值。

跟踪注册是 AR 的关键技术之一。本文以 Symbian OS 智能手机作为 AR 系统的实现平台，研究手持设备上的基于视觉标记的跟踪注册技术，通过预先在场景中设置标记，计算手机内置摄像头相对于标记的姿态（位置和方向），将虚拟物体准确地注册到标记上，达到虚实融合的目的。论文的创新成果有：

1、研究快速标记角点检测算法。标记的角点检测是标记识别以及摄像头姿态计算的首要步骤。论文研究了目前常用的标记角点检测算法，分析它们的算法性能，提出了一种利用区域生长检测标记角点的方法，实验表明与目前常用的标记角点检测算法相比，本文方法具有快速准确的特点，适合计算性能较低的手持设备。

2、解决因手持设备随用户手部无意识抖动引起跟踪注册实时性降低的问题。使用 AR 系统的平台设备时，当用户手部存在抖动情况时，摄像头姿态改变较小，若仍采用逐帧检测跟踪法，对计算性能较低的手持设备而言，则会降低跟踪注册的实时性。本文建立的基于窗口的摄像机姿态估计方法，解决了因手持设备抖动降低跟踪注册实时性的问题。本文通过大量的实验表明，通过设置大小合适的窗口，可以明显提高跟踪注册的实时性能。

3、研究手持设备上三维模型的绘制。设计实现三维模型数据解析器，将通用建模软件设计的三维模型转换成适合于手持设备渲染的模型数据格式，实现了三维模型的复用，提高了 AR 系统的开发效率。

4、设计实现了 Symbian OS 智能手机上的 AR 游戏——Bear Hitting。基于本文的研究成果，设计实现了智能手机 Nokia 6120c 上的 AR 游戏。由于键盘操作会引起手机的抖动，本文的方法可显著地提高游戏的实时响应速率。

关键词：增强现实；基于标记的跟踪注册；手持设备

Abstract

Augmented Reality (AR) technology deals with the combination of real-world and computer-generated data, where computer graphics objects are blended into real footage in real-time. AR technology applies in many fields, including military, medical treatment, education, culture and entertainment. As the increasing development of mobile technology, small and low power handheld devices, such as smart phones and PDAs, have become an ideal platform for AR. Therefore, it is meaningful for researching AR system on handheld devices.

Tracking is a most important technology in AR system. In this paper, marker tracking technology based on smart phones shipped with Symbian OS is researched. By setting a marker in real scene, camera pose according to this marker is calculated and then virtual objects could be attached on it. The main work in this paper is as followed:

Firstly, a fast algorithm of detecting corners of marker is proposed. Corners detecting is the first step for recognizing marker and calculating camera pose. By analyzing the popular methods used in corners detecting in AR system, an algorithm based on region growing for finding corners of connected regions is proposed. The experiments show that this algorithm is faster than other existed methods used aimed at finding marker's corners. Therefore it is feasible for end-users' devices with low computational resource.

Secondly, a solution for avoiding decreasing the performance of real time of AR system caused by handheld devices' joggling is presented. When user holds a phone with one hand, camera's pose is not always fixed but changes a little. For end-users' devices with low computational resource, when camera joggling, it would decrease the performance of real time of AR system if calculating the camera's pose. Therefore, an algorithm based on window estimating camera's pose is presented for this problem. The experiments show that this algorithm could improve the performance of real time greatly by setting suitable window size.

Thirdly, a model parser is presented for the aim of reusing models constructed in ordinary software.

Fourthly, an AR game, Bear Hitting, based on Symbian OS phone is developed to testify the validity of the above algorithms. The phone will be joggling because of pressing the keys. With algorithms proposed in this paper it will improve the performance of real time of AR game.

Keywords: Augmented Reality; marker tracking; handheld device

学位论文独创性声明

本人声明，所呈交的学位论文系本人在导师指导下独立完成的研究成果。文中依法引用他人的成果，均已做出明确标注或得到许可。论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

本人如违反上述声明，愿意承担由此引发的一切责任和后果。

论文作者签名：隋毅 日期：2009年5月10日

学位论文知识产权权属声明

本人在导师指导下所完成的学位论文及相关的职务作品，知识产权归属学校。学校享有以任何方式发表、复制、公开阅览、借阅以及申请专利等权利。本人离校后发表或使用学位论文或与该论文直接相关的学术论文或成果时，署名单位仍然为青岛大学。

本学位论文属于：

保密 在 年解密后适用于本声明。
不保密

(请在以上方框内打“√”)

论文作者签名：隋毅 日期：2009年5月10日

导师签名： 日期：2009年5月10日

(本声明的版权归青岛大学所有，未经许可，任何单位及任何个人不得擅自使用)

第一章 绪论

1.1 选题的背景与研究意义

1.1.1 增强现实技术概述

许多系统存在用户接口复杂的问题，用户很难使用这些系统。究其原因是由于缺少设计用户接口的能力，另一方面是由于随着系统计算能力的提高，设备和应用系统变得越来越复杂并且整合了大量的功能。因此，良好的用户接口设计对于开发高可用性应用系统来说是极其重要的。

增强现实（Augmented Reality，简称 AR）的研究目的是开发新型的人机交互接口。增强现实不是将信息孤立地呈现在设备上，而是将信息准确地注册到三维现实世界中。因此，增强现实技术模糊了真实世界和用户接口之间的关系，将二者以一种自然的方式结合，由此可以针对复杂的应用系统开发出简单实用的用户接口。

尽管第一个 AR 系统诞生于 1960s，但是 AR 从虚拟现实中分离出来，成为一个独立的研究领域是在 1990s。如今，存在两种 AR 的定义，由于缺少官方的定义，这两种定义都为研究者所接受。下面是 Azuma 对 AR 系统的定义，他认为 AR 系统必须满足以下三种特点^[1]：

- ①真实和虚拟融合；
- ②将虚拟的物体准确注册到真实世界中，注册的位置是三维的；
- ③允许实时的交互。

第一个特点是对 AR 的基本描述，是指 AR 呈现给人们的是虚实融合的景象；第二个特点是将 AR 与 Mixed-Media（混合媒体）相区别，也就是说虚拟物体必须准确地注册到真实世界中，与真实世界完美的融合，仿佛是一体的，这就要求虚拟物体的注册位置是三维的；第三个特点是指允许用户实时地与 AR 系统交互，这是 AR 系统与离线的增强系统本质的不同点，比如电影中虚拟的动画人物与真实人物同时出现的景象，就是经过长时间的离线的制作得来的，它不允许观众与其进行实时交互。

根据 Milgram 提出虚拟闭合集的描述^[2]，AR 仅仅是混合现实（MR）的一部分，MR 是将真实和虚拟世界共同呈现在同一个显示设备上。虚拟闭合集中将 AR 和 AV（Augmented Virtuality）并列。AR 显示的主体是真实世界，仅有有限的虚

拟物体混合进来；相反，AV 显示的主体则是虚拟世界，在虚拟的环境中掺杂一些真实的元素。AR 和 AV 之间的区别没有严格的定义。如图 1.1 所示

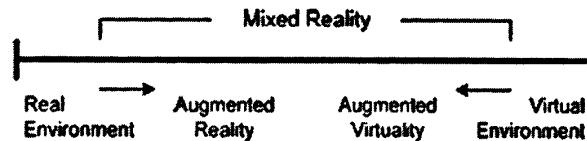


图 1.1 Milgram's 关于混合显示闭合集的描述^[2]

1.1.2 基于手持设备的增强现实系统

增强现实系统按其使用范围可分为^[3]：室内增强现实系统（Indoor AR）和室外或移动增强现实系统（Outdoor or Mobile AR）。顾名思义，Indoor AR 系统的设备一般是固定的，不允许用户四处走动，限制了用户的活动范围；Mobile AR 系统是指用户身负各种计算设备，可以在户外随意走动，而不影响系统功能的实现。由此可见，移动增强现实系统可以更好地协助用户在工作场地工作，其适用范围更广一些。传统的 Mobile AR 系统一般是由头盔显示器（Head Mounted Display，简称 HMD）、跟踪 GPS、虚拟图形生成装置组成。很显然，这些设备不仅限制了使用者的灵敏性，而且由于这些设备造价昂贵，使得它们也不为大众市场所接受。

随着各种移动技术的发展和普及，个人数字助理（PDA）和智能手机这类通用手持设备引起了研究者的注意。从普及程度上来说，这些设备具有体积小、携带方便的特点，已经融入了人们的工作和生活；从硬件条件上来说，由于具有较快的计算处理能力、内置的摄像头和彩色的显示屏幕等特点，手持设备已经成为增强现实系统良好的实现平台^[5]。Wanger 博士认为，移动增强现实系统的下一代研究将会把注意力转向小巧的、计算能力较强的手持设备上^[5]。

图 1.2 是 Wanger 在文献^[6]中所描述的移动增强现实系统的组成元素。(a)图中是传统的移动 AR 系统的组成：背负式笔记本和头盔显示器（Head Mounted Display，简称 HMD）。其优点是背负式的笔记本提供了可靠的处理能力，HMD 可以提供最佳的融合显示效果；缺点是给使用者带来了不舒适感，其推广难度也较大。(b)图中，用户使用 Tablet PC 代替了背负式笔记本和头盔显示器，Tablet PC 具有和背负式笔记本相似的处理能力，数字输入提高了交互的准确性和简便性。但是，其重量和体积对于使用者来说仍是不小的负担。(c)图中，使用 PDA 作为 AR 系统的实现平台。第一个基于 PDA 的 AR 系统出现于 2000 年^[6]，当时鉴于 PDA

处理能力的限制，采用的瘦客户端的方式，AR 的处理任务在远程服务器上实现，PDA 仅仅作为输出显示的客户端。近几年，随着移动技术的发展，智能手机的处理能力得到了极大的提高，在图(d)中，采用智能手机作为 AR 系统的实现平台成为可能。移动手机是最小巧的、并且是普及性最广的一类手持设备，较之前面提到的三种设备，手机已经完全的融合于现代人们的生活。据[8]调查报告显示，2009 年，全球将会有 8 亿手机用户，并且这些手机具有较强的处理能力，内置摄像头和彩色显示屏。研究基于智能手机的增强现实系统具有理论意义和实用价值。

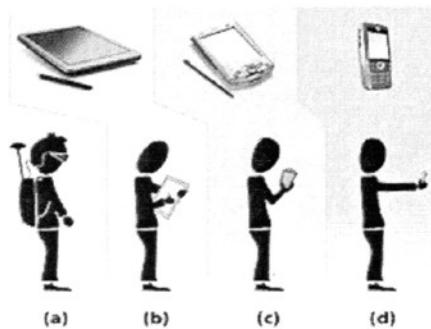


图 1.2 移动 AR 系统的实现平台^[8]

1.2 增强现实技术研究进展

1.2.1 国外研究动态

1968 年，Ivan Sutherland 教授研制出世界上第一台光学穿透式头盔显示器（See-through Head Mounted Display）^[1]，该设备可以实时地将计算机生成的简单的图形注册叠加到真实场景中。这是第一个符合 Azuma 和 Milgram 关于增强现实定义的应用，即满足虚实结合、实时交互、三维注册这三个特点。Azuma 给出了增强现实应用的领域^[1]，包括：医疗、制造与维修、机器人动作路径的规划、娱乐和军事等几个方面。1992 年，美国北卡大学（UNC at Chapel Hill）将超声波数据与病人的腹部成像集合，使外科医生可以“透视”病人的腹腔；1993 年，美国哥伦比亚大学的 Feiner 教授等人设计一个基于知识的 AR 系统（Knowledge-based Augmented Reality for Maintenance Assistance, KAEMA），该系统用于指导机械维修，可将有关技术说明叠加在激光打印机上，辅助技术人员完成维修任务；加拿大多伦多大学的 Milgram 教授等研制开发的 ARGUS 系统可以利用 AR 进行机器

人路线设计，比传统方法更易于操作。

随着功能强大的便携式计算机和笔记本电脑的推广，移动 AR 设备也有了长足的发展。哥伦比亚大学的 The Touring Machine^[9]是第一个使用了这种硬件平台的移动 AR 系统。MARS (Mobile Augmented Reality System)^[10]是该校另一个移动 AR 研究项目，它是第一个真正允许用户背负各种 AR 设备自由移动的系统。用户背负计算设备、电池、头盔显示器 HMD，可以在校园中四处行走，在已经拆除的建筑处，头盔显示器显示由计算机生成的虚拟的建筑和真实校园的融合景象。Tinmith^[50]协助用户导航的移动 AR 系统，研究人员对其进行扩展，将它用于 AR Quake^[11]户外游戏中。AR Quake^[11]是第一个将移动 AR 技术用于户外游戏的应用，用户可以通过其佩戴头盔显示器 HMD 来发现真实世界中的虚拟怪兽，游戏者可以使用特制的手枪射击怪兽。Human Pacman^[12]是另一个移动 AR 户外游戏，游戏者分为 Pacman 和 Ghost 两种角色，在户外广大的区域内，他们通过头盔显示器各自寻找真实场景中虚拟的饼干，不同角色之间可以互相射击，得分最高者即为胜利者。尽管上述移动 AR 扩大了 AR 系统的适用范围，但是由于这些 AR 系统的硬件设备沉重、复杂，给人们带来了不舒适感，因此，这类系统的推广难度较大。

随着桌面 PC 和 UMPC (Ubiquitous Mobile PC) 的普及，研究人员开始将已有的 AR 软件配置到这些通用的设备上。ARToolKit^[7]是目前比较常用的一种适用于桌面级的增强现实跟踪注册工具包，它最早是由日本教授 Hirokazu Kato 研制开发的，后来由华盛顿大学的 HIT Lab (Human Interface Technology Laboratory) 等多个研究机构共同推进完善。Kato 最早将 ARToolKit 用于实现协作式 AR (Collaborative AR) 的远程会议^[13]，该系统用户分为 AR 用户和桌面用户，两类用户均配有摄像头装置，AR 用户佩戴头盔显示器在工作场地进行工作，通过网络连接将采集到的视频图像传送给桌面用户，由桌面用户指导 AR 用户工作。由于桌面 AR 系统具有丰富的软硬件资源的支持，研究开发基于桌面级 AR 系统具有极大的便利性，但是，桌面 AR 一般只适用于实验室或固定场所，这就限制了 AR 系统的适用范围。

手持式 AR (Handheld AR) 是指由手持设备作为 AR 系统的实现平台，它成为移动 AR 的良好补充，进一步推动了 AR 的研究发展。NaviCam^[14]是第一个使用视觉穿透式的手持设备 (See Through HMD Display) 进行显示的 AR 系统，该设备通过电缆与 PC 机相连，通过识别图书上不同的彩色编码，将相应的由计算机生成的图书信息叠加到图书中，帮助使用者了解图书内容。日本的 SONY 计算机科学实验室也研究了一种和掌上电脑差不多大小的手持式显示器，利用这种显

示器构建了 TransVision^[15]协同式工作环境，用户还可以通过特制的控制器来与虚拟物体进行交互。AR Pad^[16]由一块 LCD 面板、摄像机和远程 PC 机组成，远程 PC 机配有 ARToolKit 软件包来完成跟踪注册的任务，由 LCD 面板显示最终的增强景象。除此之外，该系统也允许用户通过操纵球形输入设备来与虚拟物体进行 6DOF (six degree of freedom) 的交互。华盛顿大学 HIT Lab 研究设计的 MagicBook^[18]允许用户通过特制的手持设备浏览图书，在平面图书上显示三维立体图形，以增强用户对图书内容的理解。该系统打破了 Miligram 关于混合现实闭合集的概念，允许用户在增强现实和虚拟现实中进行切换，使得用户可选择欣赏虚实结合的画面与完全虚拟的画面。mPARD^[17]是第一个通过无线方式与 PC 机进行连接、使用手持设备进行显示的应用。上述应用都是采用特制的手持设备作为 AR 系统的显示设备，其通用性不强，这就限制了手持式 AR 的推广。

随着 PDA 和智能手机处理能力的提高和性能的丰富，研究者开始探索将 PDA 和智能手机这类通用的、普及型较高的手持设备作为手持式 AR 的实现平台。AR-PDA^[19]是第一个使用 PDA 作为 AR 系统实现平台的应用，由于受当时 PDA 处理性能的限制，PDA 仅作为瘦客户端通过无线网络传输其内置摄像头捕获到的视频流，由远程 AR 服务器完成跟踪注册的任务，之后，AR 服务器将包含有虚拟物体注册的增强视频流传回至 PDA，最后由 PDA 输出显示。这种基于瘦客户端/服务器的架构对于早期处理能力尚不成熟的 PDA 是适合的。2002 年，Wanger 将 ARToolKit 软件包移植到 PDA 上，实现了第一个由 PDA 独立完成所有 AR 处理任务的系统^[6]。Wanger 将 ARToolKit 中最为耗费计算量的部分重新进行编码，使用定点运算模拟代替浮点运算，已解决 PDA 缺少浮点运算单元 (FPU) 的问题。Wanger 所在的奥地利 Vienna University 和 Graz University of Technology 一直致力于研究手持式 AR，其开发研制的 Studierstube ES^[20]软件框架是专门针对在各种通用手持设备上进行 AR 系统开发的软件。Studierstube ES 包括：适用于智能手机和 PDA 的基于标记的跟踪库、三维图形渲染引擎和以分布式 AR 为目的的通信框架。早期的 Studierstube ES 主要针对于 Windows CE 操作系统的手持设备，近期，Studierstube ES 提供了对 Symbian OS 操作系统的接口。基于 Studierstube ES 软件框架，Wanger 所在的团队研究开发了大量的应用。AR kanji^[21]使得用户通过 PDA 识别不同的汉字卡片，并将相应的三维图形叠加于卡片之上。Invisible Train^[48]是多用户的游戏，如图 1.3 (e) 所示，游戏者可以在一个真实的木质铁路轨道上发动虚拟的火车，只有透过 PDA 显示器才能看到这些虚拟的火车，根据显示器上不同的操作指示，可以控制火车前进的方向、调整火车行驶的速度。Mr. Virtuoso^[22]

将手持式 AR 系统应用到教育领域，使用者可以通过 Mr. Virtuoso 这一虚拟的人物学习历史知识。

同 PDA 上 AR 系统类似，2004 年，Möhring 等人研制开发了第一个由智能手机完成所有 AR 任务的应用^[23]。Möhring 采用立体彩色标记进行跟踪注册，显示增强视频流的帧率是可接受的。Henrysson^[24]将 ARToolKit 工具包移植到 Symbian 手机上，实现了第一个基于 Symbian OS 的 AR 系统。最初的移植是将 ARToolKit 作为一个类添加到 Symbian 中^[49]，在 Nokia 6600 手机上获得显示帧率仅为 1fps。为了获得实时的显示帧率，Henrysson 将 ARToolKit 中最耗费计算量的代码重新编写，使用定点库代替浮点元素，提高了运算速度，可以获得较为流畅的显示画面^[24]。

近年来，基于 AR 技术的各类手持设备游戏不断涌现，如图 1.3 所示，AR 游戏的研究成为推广 AR 技术的重要途径。

Penalty Kick^[25]游戏的目的是将虚拟的足球射入印刷在海报上的球门，虚拟足球的位置由摄像头相对与海报上印刷的标记的位置决定，使用者可以旋转或是倾斜手机来入射足球，有一个虚拟的守门员将把守球门。

Mozzies Game (Mosquito Hunt)^[26] 游戏获得 2003 年最佳游戏奖，通过移动手机来瞄准射击蚊子。该游戏采用简单的像素流检测算法代替基于标记的跟踪注册方法。手机内置摄像头捕获现实世界视频流作为游戏背景，生成的虚拟蚊子注册叠加于背景之上。该游戏只测定旋转的角度，与位置远近无关。如图 1.3 (a) 所示。

Mobile Maze^[27]游戏要求使用者推动一个汽车的球通过迷宫，该游戏仅计算手机的相对于迷宫的旋转方向，不显示真实场景画面。如图 1.3 (c) 所示。

Symball^[29]游戏是一个多用户的乒乓球游戏。该游戏是基于 Symbian 操作系统的，于 2005 年 Video Processing at VTT (芬兰) 研究开发。该游戏通过检测视频流中特定物体颜色来计算手机的移动方向。球台是静态的，而球拍可以通过旋转手机进行移动，两个使用者通过蓝牙进行无线通信。

AR Tennis^[24]游戏于 2005 年 HIT lab NZ 研制开发，使用者使用标记共享一块虚拟的网球场地。手机本身做球拍，每部手机背后可以贴一张标记，使得对方手机可以检测到球拍的运动。如图 1.3 (d) 所示。

AR Soccer^[28]游戏是西门子在 2006 年世界杯期间推出的一款罚球游戏，手机屏幕的上半部分显示的是虚拟的球门和守门员，下半部分则是叠加虚拟足球后的实时捕获的视频图像。该游戏允许使用者用自己的脚来踢虚拟的足球，将虚拟的

球射入球门，通过跟踪脚的位置确定足球运动的方向。该游戏也使用像素流检测算法，可以准确地检测到移动的脚的边缘。如图 1.3 (b) 所示。

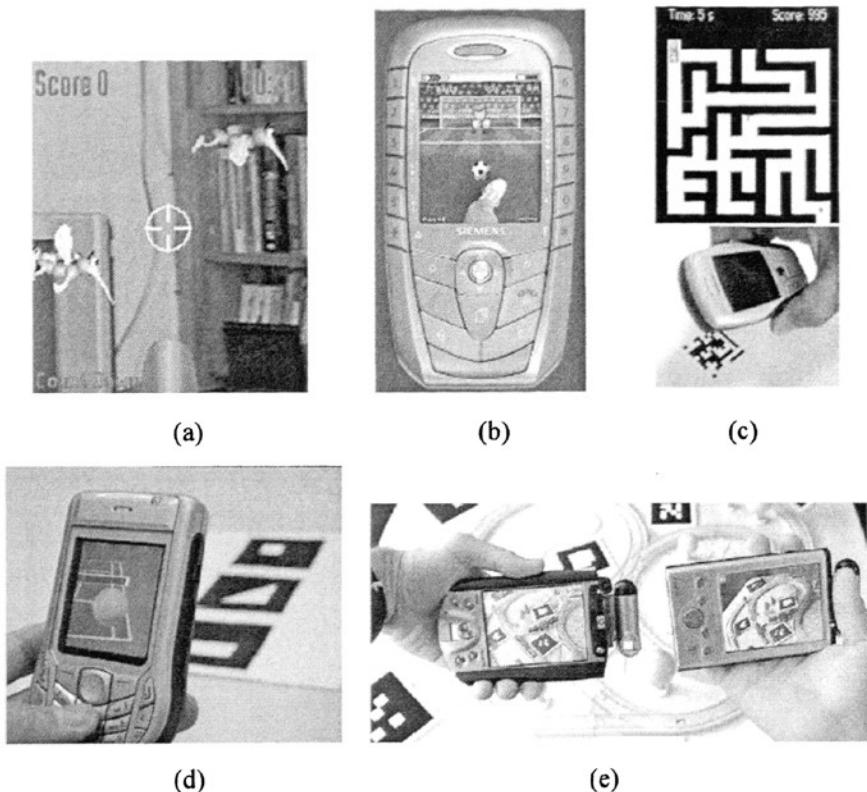


图 1.3 手持设备上的 AR 游戏：(a) Mozzies Game 游戏；(b) AR Soccer 游戏；(c) Mobile Maze 游戏；(d) AR Tennis 游戏；(e) Invisible Training 游戏

1.2.2 国内研究动态

增强现实技术作为一个新兴的研究领域，虽然在国内的研究起步较晚，但因为增强现实技术具有很大的应用和发展潜力，因此正越来越引起国内学者的关注。目前国内的不少研究机构都在致力于该领域的研究。北京理工大学已经发表了多篇基于光流^[32]、投影^[35 36]的跟踪注册技术的论文，并且在头盔显示器^[33]的研究上具有相当研究水平。其研制开发的基于 PDA 的圆明园现场重建项目是我国第一个基于通用手持设备的 AR 系统^[30]，该系统采用客户端/服务器架构，由远程服务器完成 AR 的跟踪注册任务。华中科技大学的明德烈等在注册技术上提出了全局放

射变换的方法^[37]，能够简单而快速的实现注册中的坐标转换。上海交通大学的研究 PDA 上的彩色可视标志码的设计和跟踪注册算法^[38]，取得了一定的成果。清华大学正在国家自然科学基金的支持下进行“用于内窥镜手术模拟及操作的增强现实系统”的研究。总的来说，国内在该领域的研究工作比较单一，不够系统化。

1.3 研究内容

本文的研究对象是手持设备上 AR 系统的跟踪注册技术。目前，基于 Symbian OS 的智能手机占据了市场绝大部分份额。因此，研究 Symbian 智能手机上的跟踪注册技术具有广泛的应用价值。尽管随着移动技术的发展，智能手机的计算处理能力等得到了极大的提高，但是其性能较之一般通用型计算机而言仍是不足的。因此，这就给基于这类手持设备的 AR 系统的研究提出了新的挑战。论文的研究成果有：

1、研究快速标记角点检测算法。标记的角点检测是标记识别以及计算摄像头位置和方向的首要步骤。论文研究了目前常用的标记角点检测算法，分析它们的算法性能，引入一种利用区域生长检测标记角点的方法，实验表明与目前常用的标记角点检测算法相比，本文方法具有快速准确的特点，适合计算性能较低的手持设备的情况。

2、解决因手持设备随用户手部无意识抖动引起跟踪注册实时性降低的问题。使用 AR 系统的平台设备时，当用户手部存在抖动情况时，摄像头姿态改变较小，若仍采用逐帧检测跟踪法，对计算性能较低的手持设备而言，则会降低跟踪注册的实时性。本文建立的基于窗口的摄像机姿态估计方法，解决了因手持设备抖动降低跟踪注册实时性的问题。本文通过大量的实验表明，通过设置大小合适的窗口，可以明显提高跟踪注册的实时性能。

3、研究手持设备上三维模型的绘制。设计实现三维模型数据转换器，将通用建模软件设计的三维模型转换成适合于手持设备进行渲染的模型数据格式，实现了三维模型的复用，提高了 AR 系统的开发效率。

4、设计实现了 Symbian OS 智能手机上的 AR 游戏——Bear Hitting。基于本文的研究成果，设计实现了智能手机 Nokia 6120c 上的 AR 游戏。当手机内置摄像头检测出合法标记时，会在标记上注册虚拟的小熊，用户通过手机键盘可敲击小熊。由于键盘操作会引起手机的抖动，本文的方法可显著的提高游戏的实时响应速率。

1.4 论文结构

第二章：介绍增强现实关键技术，主要包括显示技术、跟踪注册技术和目前常用的 AR 工具包和 AR 软件架构。

第三章：研究了基于视觉标记的跟踪注册技术，分析了跟踪注册的基本原理，对目前常用的标记检测识别方法进行比较，讨论了 ARToolKit 中摄像机姿态计算方法。

第四章：首先设计实现了手持设备上快速标记角点检测算法，并对算法性能进行分析；其次，设计基于窗口的摄像机姿态估计算法，解决因手持设备抖动引起的 AR 系统实时性降低的问题。

第五章：研究开发了手持设备上 AR 系统，设计实现三维模型解析器，实现 PC 机上三维模型的复用。设计实现了智能手机上的 AR 游戏——Bear Hitting 以验证第四章研究设计的算法。

结论：对全文进行总结与展望

第二章 增强现实关键技术

增强现实是一个典型的交叉学科，它的研究内容十分广泛，涉及计算机图形、图像处理、人机界面、移动计算、计算机网络、分布式计算、信息获取和信息可视化，以及新型显示器和传感器的设计等各个领域。AR 系统最终呈现给人们的是虚实融合的景象，因此，虚实融合的显示设备以及显示方式是 AR 的关键问题之一。另外，虚拟物体在真实世界中的注册位置必须是三维的，当摄像机姿态（人们的视角）发生变化时，虚拟物体的姿态也会发生相应的变化，因此，跟踪摄像机的姿态是 AR 的关键技术之一。

2.1 显示技术

增强现实系统使用多种输出设备以满足虚实融合的需要，按显示设备可以分为头盔显示器、手持设备、平面显示设备、投影设备、以及 3D 显示设备。

2.1.1 头盔显示器

1968 年，Ivan Sutherland 教授研制出世界上第一台光学透视式头盔显示器（see-through head-mounted display，简称 HMD）^[1]，该设备可以实时地将计算机生成的简单的图形叠加到真实场景中，这是第一个符合 Azuma 对增强现实定义^[1]的显示设备。

增强现实系统的头盔显示器能够看到周围的真实环境，所以叫做透视式（See-through）头盔显示器。透视式头盔显示器一般分为视频透视式（Video see-through）和光学透视式（Optical see-through）。这两类 HMD 都配有摄像机装置，佩戴者头部的位置和方向就代表了摄像机的位置和方向。Video See-through HMD 是利用摄像机对真实世界进行同步拍摄，将其送入虚拟场景工作站，在虚拟场景工作站中将虚拟场景生成器生成的虚拟物体同真实世界场景相融合，然后输出到头盔显示器显示。也就是说，使用者看到的是虚实融合的视频图像。Optical-through HMD 是在佩戴者眼前设置一块半透明的光学组合器，一方面，佩戴者可以直接看到外部世界，另一方面，该光学组合器又是部分反射的，虚拟场景生成器将虚拟物体同真实世界在人眼中融合。即，使用者可以直接看到是外部真实世界而不是经过处理之后的视频图像。

头盔显示器可以呈现给用户最佳的视觉显示效果，并且可以释放用户的双手，但是其缺点也是显而易见的，其体积较大、佩戴不方便、长时间佩戴会引起不适

感，对于使用者来说较难接受。

2.1.2 手持设备显示

手持设备（Handheld display）显示器克服了头盔显示器的不足，它允许用户手持显示设备，消除了佩戴 HMD 的不舒适感，但同时也减弱了沉浸感。NaviCam^[14]是第一个使用视觉穿透式方式的手持设备（See Through HMD Display）进行显示的 AR 系统，它使用一块 LCD 显示屏显示虚实结合的场景。MaigicBook^[18]是华盛顿大学 HIT Lab 研制开发的手持显示设备，用户通过该设备既可以体验虚实结合的场景，也可以沉浸其中，感受完全虚拟的场景。

随着移动技术的发展，各种通用手持设备，如 PDA、智能手机等，成为增强现实系统良好的实现平台^[5]。目前绝大多数智能手机、PDA 具有内置摄像头和彩色显示屏，符合增强现实系统的显示要求。虽然这类设备的显示屏幕较小，但是由于其普及型较高，智能手机成为下一代移动 AR 系统平台的发展方向^[5]。

2.1.3 其它几种显示设备

1、视网膜显示^[39]

Microvision 公司提供了商用显示系统 Nomad，这个系统可以将虚拟图像直接投射到视网膜上，这种方法在一定程度上减少了显示设备的体积，提高了输入光线的强度。Microoptical 的 eyeglass 系列产品，将一个小彩色显示屏直接安装在普通眼睛上，这种设备仅重几十克，但是显示区域过小，其应用受到限制。

2、平面显示设备

最常用的个人计算机可以作为 AR 系统的显示设备，这类设备造价较低，满足了大量的桌面级用户的需求。

3、投影设备^[39]

投影设备能将图像投影到大范围的环境中。与 HMD 显示相比，投影设备适合室内 AR 应用环境，其生成图像的角度不随使用者的视角改变而变化。同时，投影设备能够直接将虚拟物体投影到真实世界的相应位置，改变真实物体的表面纹理和真实场景的光照效果。

2.2 跟踪注册技术

AR 系统要求实现虚拟物体和现实场景的无缝融合，即，需要将虚拟物体与真实世界在三维空间位置中进行“配准”，使人们感觉虚拟物体仿佛是真实存在于真

实场景中的。这个“配准”的过程常称为注册（registration）。要实现虚拟物体在真实场景中的精确注册则需要从当前场景中获取空间数据。虚拟物体所处位置和观察者位置是相对的，这意味着 AR 系统必须实时地检测观察者（摄像机）在真实场景中的位置和方向，以帮助 AR 系统获得观察者（摄像机）的视角，这个过程成为跟踪（Tracking）。在 AR 系统中，跟踪直接影响到虚拟物体的注册，因此，跟踪技术是 AR 的关键技术。在 AR 中经常使用如下跟踪技术^[39]：

- 1) 视频检测：该方法是 AR 系统中常采用的跟踪注册方法，又称为基于计算机视觉的跟踪注册方法。其执行流程为：①通过摄像机获取真实场景信息；②对真实场景信息进行分析，计算摄像机在真实场景中的姿态；③生成虚拟物体；④将虚拟物体注册到真实场景中，输出显示增强视频。在视频检测方法中，为了确定摄像机相对于真实场景的位置，通常需要设置参考点，因此，该技术分为：基于标记的视频检测技术和基于自然特征点的视频检测技术。前者要求在场景中预先设置人造标记，而基于自然特征点的视频检测方法则通过检测场景中的自然特征，比如，点、直线或是曲线，以此来确定摄像机的相对位置。
- 2) 光学系统：通过使用 CCD 传感器测量各种目标物体和基准上安装的 LED 发出的光线来测量角度，通过这些角度计算目标的方向和距离。
- 3) GPS：用于户外 AR 系统中跟踪和确定用户的地理位置。
- 4) 超声波：原理与 GPS 类似，利用测量接收与三个已知超声波源的距离判断使用者的位置。
- 5) 惯性装置：通过惯性原理来测定使用者的运动加速度。
- 6) 陀螺仪：用于测定使用者头部转动的角度，来判定如何转换视场中虚拟物体的坐标。
- 7) 磁场：通过感应线圈的电流强弱来判断用户与人造磁场中心的距离，或利用地球磁场判断方向，如电子罗盘。
- 8) 机械装置：利用接卸装置的节点之间的长度和角度来定位各个节点。

目前已有的大型 AR 系统均综合采用以上跟踪注册方法。比如，The Touring Machine^[9]使用了 GPS、电子罗盘组成的混合跟踪定位系统，来判断用户在校园中的位置、视野的方向和头盔的角度；AR Quake^[11]游戏使用 GPS、电子罗盘和视频检测相结合的方法，利用 GPS 和电子罗盘进行大范围粗略的定位，之后采用视频检测方法进行精确跟踪。

2.2.1 基于标记的视频检测技术

基于标记的视频跟踪注册技术需要预先在真实场景中设置人造标记，其执行流程为：①AR 系统对捕获到的场景信息进行标记检测、识别，以此确定标记的 ID；②计算摄像机相对于该标记的姿态；③生成标记 ID 对应的虚拟物体；④根据摄像机的姿态，注册虚拟物体，最后输出合成的画面。标记设计通常为黑白平板样式，这样可以较快的从复杂环境中检测出是否存在标记，如图 2.1 所示。

标记的设计与识别方法是研究基于标记的 AR 系统必须要面对的问题。按照标记设计方法的不同，标记的识别分为：①基于图像模板匹配的标记识别；②基于编码特征的标记识别。前者的标记设计可具有可读性，如图 2.1 (a) 所示，通过在系统中设定标准标记模板，将检测到的候选标记与标准模板进行逐一匹配，确定最佳的候选匹配，该方法简单易行，但当需要识别多个标记时，搜索匹配的时间较长，计算量较大，影响 AR 系统的实时性；后者按照二进制编码的方式设计标记，提高了可识别标记的数量。通过对标记内部图形进行解码识别标记，并采用循环冗余码 CRC 等编码校验技术提高了标记识别的准确性，具有一定的鲁棒性。

桌面级 AR 系统的开发工具包 ARToolKit^[4]采用基于图像模板匹配的方式设计识别标记，如图 2.1 (a) 所示。通常，此类标记的设计具有可读性，标记本身携带一定的信息量。该方法是目前 AR 系统较多采用的方法之一。

加拿大国家研究委员会（National Research Council Canada）设计实现 ARTag^[44]，如图 2.1 (b) 所示。ARTag 采用二进制编码的方式扩展了可识别标记的数量，标记库容量达到 2^{10} 种，并且利用编码校验技术提高了标记识别的准确性，但是其算法较为复杂，对硬件资源要求较高，不适应于手持设备。

瑞士苏黎世理工学院 Rohs 教授等研究手持设备上的 AR 技术，设计实现了 Visual Code^[43]标记，如图 2.1 (c) 所示。它采用类似二维条码的方式设计标记，扩大了可识别标记的数量，但是其只能提供三自由度的摄像机姿态信息，虚拟物体的注册位置不够精确。

奥地利的计算机图形与视觉研究所与 Graz University of Technology 联合开发 ARToolKitPlus^[51]，如图 2.1 (d) 所示。该开发工具包的目标是研发各类手持设备上的 AR 系统，其实现平台包括 UMPC、PDA、智能手机等。ARToolKitPlus 中采用 BHC 编码技术设计标记，可识别标记数量达到 2^{12} 种，其标记识别的效率较 ARTag 高。但是，当标记被部分遮挡时，该类标记的识别方法将会失效。

文献[41]设计一种新的标记系统，通过在标记中引入代码区与直观的图像信息，提高了标记识别的准确性。文献[38]研究手持设备 PDA 上彩色可视码的设计与识别。

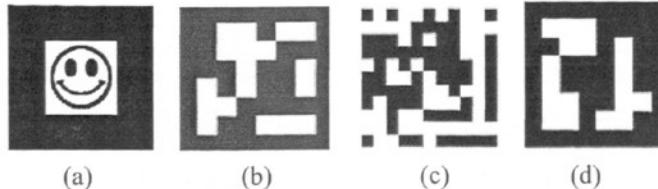


图 2.1 黑白平板标记：(a) ARToolKit 中标记；(b) ARTag 标记；
(c) Visual Code (可视码)；(d) ARToolKitPlus 中标记

2.2.2 基于自然特征的视频检测技术

基于标记的视频跟踪注册方法需要在场景中预先设置标记，在户外时，很难做到在每个场景中都设置标记，另外，对于活动范围较大的 AR 系统，需要大尺寸的标记才能满足要求，这也给用户带来了诸多不便。因此，研究基于无标记的跟踪注册技术可以扩大增强现实系统的适应范围。为了能够确定摄像机在真实场景中的姿态，可以将真实环境中的某些自然特征作为参照物，比如，点、直线或是曲线，通过特征提取和特征匹配实现三维位置的注册。该种方法计算量较大，对实时性要求较高的增强现实系统来说是不适合的。

现阶段，国内外学者对该技术的研究主要集中于通用计算机，手持设备、特别是通用手持设备上的研究则较少。2008 年，Daniel 博士首次将基于自然特征点的跟踪注册方法应用于智能手机上^[45]。基于光流法（optical flow）的自然特征跟踪是目前手持设备上 AR 系统较常用的一种跟踪注册方法。Siemens SX-1 上的 Mosquito Hunt^[26]游戏采用该方法实现了无标记的跟踪注册，这之后，出现了许多采用该方法的 AR 手机游戏。Wang 等研究设计的开源的基于像素流的跟踪库—TinyMotion^[46]可协助开发者进行这方面的应用开发。

2.3 AR 系统开发工具

2.3.1 AR 工具包

1、ARToolKit^[4]

为降低 AR 应用开发的复杂性，许多研究组织陆续提出用于构建 AR 系统的开发工具包，这些工具包中包含了 AR 开发中涉及的模式识别、坐标转换以及视频合成等功能。ARToolKit 是一套开放源码的工具包，最初由日本大阪大学的 Hirokazu Katob 博士开发，后来由华盛顿大学 HIT Lab 和新西兰的坎特伯雷大学 HIT Lab NZ 共同资助。ARToolKit 已经成为目前使用最广泛的桌面级 AR 系统开发包，现有的许多 AR 应用都是基于 ARToolKit 实现或利用类似的改进版本进行开发的。ARToolKit 采用基于标记的视频检测技术进行跟踪注册，使用 OpenGL 三维图形渲染库生成 3D 虚拟图形，输出显示虚实融合的视频流中。ARToolKit 提供 Linux、Mac 和 Windows 系统下的不同版本，方便研究者使用。

2、ARLib^[52]

ARLib 是英国 Bournemouth University 大学于 2005 年研究开发的 AR 工具包，其主要功能与 ARToolKit 类似，对 ARToolKit 中的标记检测和识别进行了改进，其标记内部由黑色圆点按照一定规则排列，并且，采用动态阈值的方法进行图像分割，解决了当光照条件发生变化时，标记准确识别的问题。

3、ARToolKitPlus^[51]

ARToolKitPlus 是 ARToolKit 的改进版本，针对手持设备上 AR 系统的设计实现，由奥地利的 Graz University of Technology 的 Handheld Augmented Reality 项目研究开发。ARToolKitPlus 增加了内存管理、不同层次代码的优化功能；采用基于编码特征的标记设计、识别方法，采用启发式的动态阈值方法，解决了因光晕情况下标记无法检测的问题。截止到 2006 年 6 月，该工具包已不再进行更新，项目组将其添加入 Studierstube ES^[5]软件框架中，更名为 StbTracker。

2.3.2 AR 软件框架

Studierstube ES^[5]是一套基于手持设备的 AR 系统开发的软件框架，由奥地利的 Graz of University Technology 研制开发，它是目前唯一一套针对手持设备上 AR 系统开发的商业软件。StbES 主要针对 Windows CE 手持设备，近期，添加了 Symbian OS 接口。StbES 包括以下功能组件：①StbTracker 组件完成标记的跟踪定位；②Muddleware 组件实现多用户系统的功能，支持多用户系统的开发；③设计开发了 3D 图形渲染库 FPK 完成三维动画渲染。该项目组基于 Studierstube ES 设计开发了 Map Tracking、Cows and Aliens 和 MARQ 等应用^[53]。

2.4 小结

本章讨论了增强现实系统的关键技术，主要涉及显示技术与跟踪注册技术，并且，对目前常用的增强现实系统的开发工具包和开发平台进行介绍。在下一章将对基于视觉标记的跟踪注册技术进行研究，讨论跟踪注册的数学原理，标记检测、识别方法，与 ARToolKit 中摄像机姿态计算的过程。

第三章 基于视觉标记的跟踪注册技术

基于视觉标记的跟踪注册技术的基本思想为：预先在真实场景中设置标记，AR 系统对摄像机捕获的视频图像进行分析，检测识别图像中是否存在合法的标记，计算摄像机在真实世界坐标系统中相对于标记的姿态，将虚拟物体以三维“对齐”的方式注册到标记上，输出虚实融合的场景（如图 3.1 所示）。该方法不需要复杂、昂贵的跟踪设备，只需简单的摄像头即可；并且，通过设置标记，确定了摄像机在真实世界坐标系统中姿态，与基于自然特征的跟踪注册技术相比具有较大的效率优势。因此，目前大多数 AR 系统采用基于视觉标记的跟踪注册技术。基于视觉标记的跟踪注册技术主要涉及到标记检测、标记识别、以及计算摄像机姿态技术。3.1 小节给出了跟踪注册的基本数学原理，3.2 小节、3.3 小节分别讨论了标记检测技术与标记识别技术，3.4 小节给出了 ARToolKit 工具包中计算摄像机姿态的过程。

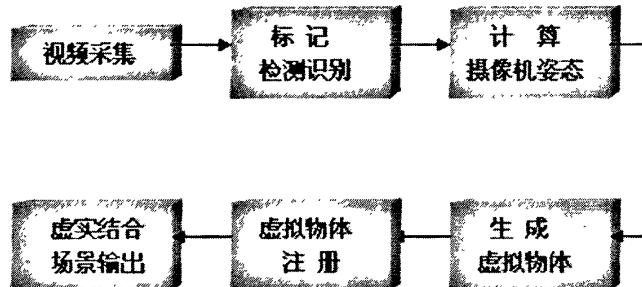


图 3.1 基于视觉标记的 AR 系统流程

3.1 基本原理

AR 的目的是将虚拟物体准确地注册到真实场景中，当摄像机姿态发生变化时，虚拟物体的注册位置也要发生相应的变化。由此可见，AR 系统涉及到虚拟物体坐标系、真实场景坐标系、摄像机坐标系、2D 成像平面坐标这 4 个坐标之间的转换。跟踪是指实时地计算摄像机的姿态，即，获得摄像机坐标系与真实世界坐标系之间的转换关系；注册是指实时地获得虚拟物体坐标系与 2D 成像平面

坐标系之间的转换关系。假设摄像机坐标空间为 $X_c Y_c Z_c$ ，真实场景坐标空间为 $X_w Y_w Z_w$ ，虚拟物体坐标空间为 $X_v Y_v Z_v$ 。AR 系统要实现的注册即是获得 $X_v Y_v Z_v$ 到 $X_w Y_w Z_w$ 的转换关系。

3.1.1 摄像机与真实场景坐标系之间的转换

摄像机坐标空间 $X_c Y_c Z_c$ 到真实场景坐标空间 $X_w Y_w Z_w$ 的坐标转换是确定摄像机在真实场景位置和方向，即摄像机的外部姿态。公式 3.1 表示了摄像机坐标空间中的某点 (X_c, Y_c, Z_c) 映射到真实场景空间 (X_w, Y_w, Z_w) 的关系：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} V_{3 \times 3} & W_{3 \times 1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = T_{cm} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad 3 - (1)$$

公式 3 - (1) 反映了摄像机与真实场景空间之间的欧几里德变换关系：对于真实空间坐标系 $X_w Y_w Z_w$ 点 P (X_w, Y_w, Z_w) ，其在摄像机坐标系 $X_c Y_c Z_c$ 有另外一个坐标值 (X_c, Y_c, Z_c) ，二者之间的变化关系可用一个 4×4 变换矩阵 T_{cm} 表示，其中 $V_{3 \times 3}$ 是一个 3×3 矩阵，反映摄像机相对于真实空间坐标系的 3 个旋转分量，表达了坐标轴的 3 个基本旋转，即绕 X_w , Y_w , Z_w 轴的旋转。 $W_{3 \times 1}$ 是一个 3×1 矩阵，反映了摄像机相对于真实空间坐标系的 3 个平移分量。由于 $V_{3 \times 3}$, $W_{3 \times 1}$ 完全由摄像机相对于真实空间坐标系的位置和方向确定，因此称 T_{cm} 为摄像机外部参数（姿态），跟踪的结果就是获得 T_{cm} 。

由于 (X_w, Y_w, Z_w) 是可以事先确定的，根据 T_{cm} ，则可以通过公式 3 - (1) 计算得到其对应的摄像机空间中的位置。

3.1.2 摄像机与 2D 成像平面之间的转换

摄像机与 2D 成像平面的转换关系可看作是摄像机成像模型，图 3.2 为摄像机的小孔成像模型。摄像机空间坐标系的原点为 O ，2D 成像平面坐标系的原点 o 位于摄像机坐标空间的 Z_c 轴上， f 为两个原点之间的距离，称其为摄像机的焦距。

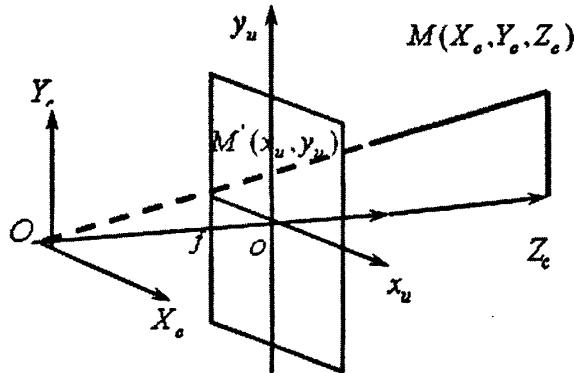


图 3.2 摄像机小孔成像原理图

对于针孔摄像机模型，假设摄像机空间中某点 $M(X_c, Y_c, Z_c)$ 在 2D 成像平面上的投影为 $M'(x_u, y_u)$ ， M' 是原点 O 与 M 的联系 OM 同 2D 成像平面的交点，由相似比例关系式得出：

$$x_u = \frac{X_c}{Z_c} f \quad y_u = \frac{Y_c}{Z_c} f \quad 3-(2)$$

上述透视投影关系用齐次坐标与矩阵表示为：

$$\begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad 3-(3)$$

由于上式中图像坐标 x_u, y_u 是实数，而计算机对图像进行处理时都是以像素为单位，所以还应把图像中点的实际位置转换成像素点位置。假设 (u, v) 是 M' 的像素坐标，每一像素在 x_u 轴和 y_u 轴方向上物理尺寸为 dx 、 dy ， (u_0, v_0) 表示图像像素的中心点坐标，于是 (x_u, y_u) 和 (u, v) 关系可表示为下式：

$$u = \frac{x_u}{dx} + u_0 \quad v = \frac{y_u}{dy} + v_0 \quad 3 - (4)$$

公式 3 - (4) 用矩阵形式表示为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} \quad 3 - (5)$$

$$\begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = \begin{bmatrix} dx & 0 & -u_0 dx \\ 0 & dy & -v_0 dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad 3 - (6)$$

将公式 3 - (6) 代入公式 3 - (3) 得：

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

通常将上式表示成以下形式：

$$\begin{bmatrix} hu \\ hv \\ h \\ 1 \end{bmatrix} = P \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}, \quad \text{其中 } P = \begin{bmatrix} s_x f & 0 & x_0 & 0 \\ 0 & s_y f & y_0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad 3 - (7)$$

P 表示摄像机内部参数矩阵, 其中, s_x 表示 x 轴方向上的缩放因子(pixel/mm), s_y 表示 y 轴方向上的缩放因子(pixel/mm), f 表示摄像机焦距, (u_0, v_0) 表示图像像素的中心点坐标。

3.1.3 虚拟物体与真实场景坐标空间的转换

虚拟空间坐标系 X_v, Y_v, Z_v 中三维虚拟物体上某点 (X_v, Y_v, Z_v) 到真实场景空间坐标系 X_w, Y_w, Z_w 中的几何描述 (X_w, Y_w, Z_w) 的转换关系表示为:

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = A \begin{bmatrix} X_v \\ Y_v \\ Z_v \\ 1 \end{bmatrix} \quad 3 - (8)$$

公式 3 - (8) 中, A 为已知的变换矩阵。

将公式 3 - (1)、公式 3 - (7) 和公式 3 - (8) 组合, 则可获得虚拟物体上的某点到 2D 成像平面上的相对应点的转化关系式, P 由摄像机定标得到:

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} Tcm \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = P \cdot Tcm \cdot A \cdot \begin{bmatrix} X_v \\ Y_v \\ Z_v \\ 1 \end{bmatrix} \quad 3 - (9)$$

为了方便起见, 通常设虚拟物体与真实场景坐标空间的转换矩阵 A 为单位矩

阵，因此，由公式 3-(9) 可见，只要确定了摄像机内部参数 P 和摄像机姿态 T_{cm} ，就可计算得到虚拟空间中的任意一点 (X_v, Y_v, Z_v) 注册到摄像机的 2D 成像平面上的位置，完成虚拟物体的注册。

由第二章跟踪注册技术的讨论可知，为了确定摄像机在真实场景中的姿态，通常在真实场景中设置标记，以此作为参考点，计算摄像机相对于标记的姿态，这种基于标记的跟踪注册技术与基于自然特征的跟踪注册技术相比具有效率的优势，是目前绝大多数 AR 系统的选择。在下面的小节中，将主要讨论几种开发工具包中使用的标记检测、识别方法并对它们的性能进行分析，着重讨论 ARToolKit 中摄像机跟踪的方法。

3.2 标记检测技术

标记检测识别的目的是判断采集的场景图像中是否存在合法的标记，标记的检测是处理的第一步。根据标记的特征，采用不同的检测方法，本文不对彩色和立体标记进行讨论，仅研究常用的黑白平板标记，如图 3.3 所示，此类标记的外部是黑色粗边框，边框内部设置不同图形确定标记的 ID。标记检测的目的是获得标记的特征，即，标记的四个角点，角点的位置是识别标记、跟踪摄像机姿态的前提。

由于标记具有明显的黑白特征，因此，首先对图像进行阈值分割是常用的标记检测预处理方法，快速、有效的阈值分割技术可以提高标记检测的性能。下面对图像阈值分割及标记特征提取技术进行讨论，分析几种常用工具包的标记检测性能。

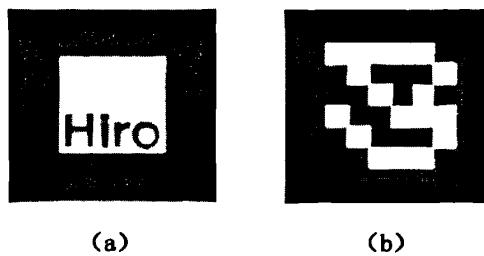


图 3.3 黑白平板标记：(a) ARToolKit 中标记 (b) ARTag 标记

3.2.1 图像阈值分割

图像阈值分割是将图像按照某一阈值分成背景和物体两部分。根据选取阈值的不同方法，主要分为全局阈值分割、局部阈值分割和动态阈值分割。全局阈值分割最为简单快速，是指设置一个固定阈值，整幅图像的像素值与其比较，若像素值小于该阈值则置为背景，否则为物体。局部阈值是指存在多个阈值，根据像素所处图像的位置，与相应的阈值进行比较，若小于则为物体，大于则为背景。全局阈值分割和局部阈值分割采用固定的阈值，动态阈值则根据外部条件（如光照等）进行变化，动态的设置阈值。

3.2.1.1 全局阈值分割法

在 ARToolKit^[4]工具包中，使用固定的全局阈值对整幅图像进行分割，该阈值可根据不同系统的需要进行设置，默认值为 100。该方法具有快速的特点，适合于对实时性要求较高的 AR 系统。

3.2.1.2 动态阈值法

文献[7]对 ARToolKit 采用的全局静态阈值的方法进行改进，解决方案为：首先使用全局固定阈值的方法，对图像进行二值化，当检测识别出场景中存在合法标记时，记录当前标记所在图像区域的局部平均阈值，以此作为下一帧中图像分割的阈值。由此可见，该方法能够解决当光照条件突然发生变化时标记无法检测的问题，但是具有一定的局限性，必须以检测到标记为前提。

ARToolKitPlus^[5]采用启发式动态阈值的方法，首先设置一个固定阈值，当检测识别到合法标记时，记录当前标记所在图像区域的局部半局阈值，当检测不到标记时，采用“猜”的方式上下调节阈值，对图像进行阈值分割。

3.2.1.3 局部动态阈值分割法

文献[54]中采用局部动态阈值分割图像，对光强发生变化、出现光晕情况，具有较好的阈值分割结果，但是需要两遍扫描图像，效率有所下降。其基本原理为：指定一个窗口，在整幅图像上移动该窗口，以当前像素为中心，计算像素周围窗口大小区域内像素的平均值，将该值与当前像素进行比较，以此来确定当前像素是背景或是物体。由于需要计算多个相互重叠的窗口的像素和，文献[54]采用了积分图像的方法，第一遍扫描源图像生成积分图像，根据积分图像，再对源图像进行扫描对其二值化。

设 $f(x, y)$ 表示源图像, $I(x, y)$ 表示积分图像, $I(x, y)$ 的意义为源图像像素 (x, y) 的左、上以及左上像素的和。当顺序扫描源图像时, 按照公式 3 - (10), 可以在线性时间内获得积分图像:

$$I(x, y) = f(x, y) + I(x-1, y) + I(x, y-1) - I(x-1, y-1) \quad 3 - (10)$$

图 3.4 (a)举例说明了计算积分图像。一旦获得了积分图像, 任意矩形 (左上角 (x_1, y_1) , 右下角 (x_2, y_2)) 的区域和可以在常量级时间内计算出来, 如公式 3 - (11) 所示。图 3.4 (b)说明了计算矩形 D 的区域和, 等同于计算多个区域的和 $(A+B+C+D)-(A+B)-(A+C)+A$ 。

$$\sum_{x_1}^{x_2} \sum_{y_1}^{y_2} f(x, y) = I(x_2, y_2) - I(x_2, y_1-1) - I(x_1-1, y_2) + I(x_1-1, y_1-1) \quad 3 - (11)$$

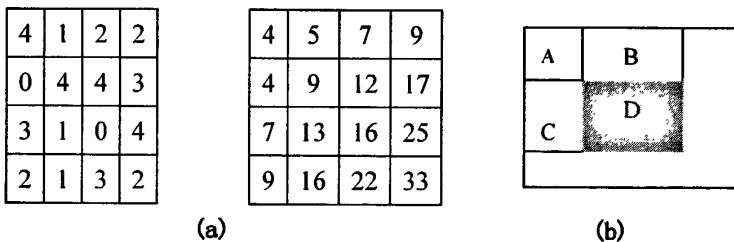


图 3.4 积分图像: (a)左边为源图像, 右边为根据公式 3- (10) 计算得到的积分图像; (b)图为使用积分图像计算灰色窗口区域的像素和

实际应用时, 设置一个 $t (0 < t < 1)$ 值, 当窗口中心像素小于窗口区域像素平均值的 $t \times 100\%$ 时, 则将该像素设为物体, 否则设为背景。

3.2.2 标记特征提取

图像阈值分割后, 需要提取标记的特征, 以便进行标记识别, 确定标记的ID。标记黑色边框的外部四个角点是标记的特征, 标记特征提取的结果就是获得标记

的四个角点的位置坐标。

3.2.2.1 直线相交求交点法

直线特征具有稳定的特点，其检测方法较多，如最小二乘直线拟合法、Hough 变换法等，角点的位置可通过直线相交计算得到。在 ARToolKit 工具包中，检测图像中的黑色连通区域，用最小二乘法拟合直线，再根据直线相交求出标记的四个角点^[13]。

3.2.2.2 最远距离求角点法

在 ARToolKitPlus 中使用最远距离的方法检测角点^[15]。其执行流程为（如图 3.5 所示）：①检测黑色的连通区域，排除面积较小的；②随机的选取连通区域边框上一点 x ，计算该边框上距离 x 最远的点，即为角点，记为 c_0 ；③计算该连通区域的中心点 $center$ ；④连接 c_0 和 $center$ ，计算边框上距离该连接直线最远的两点，即可确定另外两个角点，记为 c_1 ， c_2 ；⑤最后确定第四个角点，连接 c_1 c_2 ，计算距离边框上距离该直线最远的新点，记为 c_3 。该方法利用最远距离的方法直接检测角点，但是需要多次计算最远距离，计算量稍大。

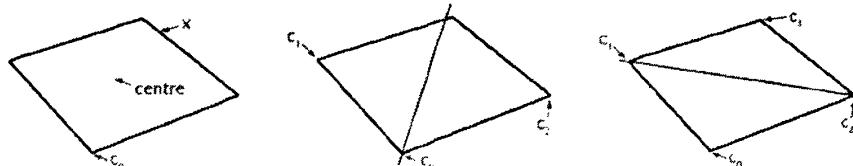


图 3.5 ARToolKitPlus 中角点检测的方法^[15]

以上两种检测标记角点的方法较为复杂，为了提高标记角点检测的效率，论文第四章设计了一种直接检测角点的方法，根据标记具有粗的黑色边框的特性，使用区域生长的方法，直接检测出标记的四个角点。

3.3 标记识别技术

标记识别的目的是为了获得标记的 ID，由此得到相对应需要渲染的虚拟物体。由于采集到的场景图像根据识别方法不同，分为基于图像匹配的标记识别方

法和基于编码特征的标记识别方法两类。

3.3.1 基于图像匹配的标记识别法^[13]

基于图像匹配的方法是设定标准标记模板，将检测到的候选标记与标准模板进行逐一匹配，确定最佳的候选匹配。由于摄像机采集场景图像时的视角是随机的，候选标记的外形轮廓发生了一定的变化，比如，缩放、旋转等，因此需要首先对标记进行校正，之后才能将其与标准标记模板进行匹配。具体实现过程如下：

1、计算单应性矩阵 H

单应性关系是指某一平面上的点映射到另一平面的线性对应关系。在标记检测过程中，可以得到标记的四个角点在成像平面上的坐标位置 $(x_c, y_c), i = 1, 2, 3, 4$ ，因为标记是预先设定的，因此，标记四个角点在真实世界的位置 $(x_m, y_m), i = 1, 2, 3, 4$ 是已知的，通常设标记的中心为原点，标记的边长为 10cm 左右，由此可以得到它们之间的单应性关系，如公式 3 - (12) 所示，列方程组可解得单应性矩阵 H 。

$$\begin{bmatrix} hx_c \\ hy_c \\ 1 \end{bmatrix} = H \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & 1 \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ 1 \end{bmatrix} \quad 3 - (12)$$

2、根据 H 将标记内部图形进行校正（Unwrap）

为了使摄像机采集到的图像与标准模板进行匹配，需要将采集的图像“正”过来，利用单应性矩阵 H 将标记内部进行映射，可达到以上目的。设标记内部某点 (x_c, y_c) ，校正之后的相对应的点位 (x_m', y_m') ，由公式 3 - (12) 计算得到的 H ，可得如下关系式：

$$H^{-1} \begin{bmatrix} hx_c \\ hy_c \\ 1 \end{bmatrix} = \begin{bmatrix} x_m' \\ y_m' \\ 1 \end{bmatrix} \quad 3 - (13)$$

3、将归一化后的图像与标准模板图像进行匹配

在 AR 系统中，事先设置在三种不同光照条件下的标记的灰度图像标准模板，

将归一化后的图像与这些模板文件按照四个不同方向进行逐一匹配，匹配次数为 12 次（如图 3.6 所示）。匹配接近的程度可借助向量的范数来描述，如欧几里得距离。设置可信度计算候选标记为合法标记的置信度 $cf(0 < cf < 1)$ ，当小于 cf 时，标记为非法，否则获得标记的 ID。

当需要识别 N 个标记时，共需匹配 $N \times 12$ 次，匹配的时间较长，该方法计算量较大。



图 3.6 在三种不同光照条件下，记录标准标记四个不同方向的灰度模板文件

3.3.2 基于编码特征的标记识别法^[43,44,51]

基于编码特征的标记识别方法的思想是：采用对标记内部图形进行二进制编码的方式设计标记，标记的 ID 由编码决定。标记进行识别时，则通过对标记内部进行解码得到相应的 ID，并且，利用编码校正技术，可以解决标记错误识别的问题，其可靠性较高。与基于图像匹配的标记识别方法相比，基于编码特征的标记识别方法可以降低误识别率，并且可识别的标记数量较大，识别速度较快。

3.4 摄像机姿态计算^[13]

当确定候选标记的 ID 后，需要计算摄像机相对于标记的姿态，下面将讨论 ARToolKit 中计算摄像机姿态的方法。

标记在真实世界坐标系统中的位置设定为：标记的中心为坐标系的原点，标记位于坐标系的 $X Y$ 平面上，标记的一对平行边平行于坐标系的 X 轴，另一对平行边平行于 Y 轴。

由标记的检测可得到一对平行边在 2D 成像平面上的直线方程：

$$a_1u + b_1v + c_1 = 0 , \quad a_2u + b_2v + c_2 = 0 \quad 3 - (14)$$

将公式 3-(7) 代入上述直线方程, 可得:

$$\begin{aligned} a_1 p_{11} X_c + (a_1 p_{12} + b_1 p_{22}) Y_c + (a_1 p_{13} + b_1 p_{23} + c_1) Z_c &= 0 \\ a_2 p_{11} X_c + (a_2 p_{12} + b_2 p_{22}) Y_c + (a_2 p_{13} + b_2 p_{23} + c_2) Z_c &= 0 \end{aligned} \quad 3-(15)$$

其中, p_{ij} ($i, j = 1, 2, 3$) 表示摄像机内部参数矩阵 P 的元素。

这两个平面的法向量分别是 n_1 、 n_2 , $n_1 \times n_2$ 得到的是这一对平行边在摄像机坐标系下的方向向量 u_1 ; 同理, 可以得到另一对平行边在摄像机坐标系下的方向向量 u_2 。理论上, u_1 , u_2 应该垂直, 但实际上是有误差的, 所以另设两个垂直的方向向量 v_1 , v_2 (如图 3.7 所示)。

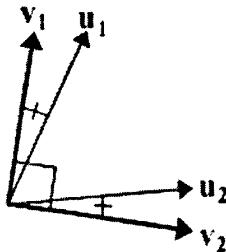


图 3.7 v_1, v_2 由 u_1, u_2 计算得来

$v_3 = v_1 \times v_2$, 所以标记相对于摄像机的旋转矩阵 $V = (v_1', v_2', v_3')$, 又由于标记的四个角点在真实世界坐标系统中的坐标位置和它们对应的在 2D 成像平面坐标都已知, 通过公式 3.7 和公式 3.1 就可计算出标记在摄像机坐标系统的偏移分量 W_{3d} 。至此, 求出 T_{cm} 。

3.5 小结

本章讨论了增强现实系统中基于视觉标记的跟踪注册技术, 首先详细地描述了跟踪注册的数学原理, 然后, 论述了标记检测和标记识别技术, 分别研究了

ARToolKit、ARToolKitPlus 中采用的标记检测方法，对目前两类标记识别方法：基于图像匹配和基于编码特征的标记识别方法进行讨论，最后，描述了 ARToolKit 中摄像机姿态标定的方法。通过本章的研究，为本论文研究手持设备上的增强现实系统奠定了理论基础。

第四章 手持设备上基于视觉标记的跟踪注册技术

随着各种移动技术的发展和普及，个人数字助理（PDA）与智能手机这类手持设备引起了研究者的注意。从普及程度上来说，这些设备具有体积小、携带方便的特点，已经融入了人们的工作和生活；从硬件条件上来说，由于具有较快的计算处理能力、内置的摄像头和彩色的显示屏幕等特点，手持设备已经成为增强现实系统良好的实现平台^[5]。Wanger 博士认为，移动增强现实系统的下一代研究将会把注意力转向小巧的、计算能力较强的手持设备上^[5]。

尽管手持设备的硬件资源不断丰富，但是与通用计算机相比仍有一定差距，因此，手持设备上的 AR 系统对跟踪注册的效率有较高的要求。本章针对手持设备硬件资源有限的问题，研究手持设备上基于视觉标记的跟踪注册技术：①研究标记检测算法，设计一种直接检测标记角点的算法（Corners Finding Directly，简称 CFD），提高标记检测的速度；②针对手持设备抖动影响跟踪注册实时性的问题，设计一种基于窗口的摄像机姿态估计算法，提高系统跟踪注册的实时性。

本文算法实验环境为：选用 Symbian OS 智能手机 Nokia 6120c 作为系统硬件平台，处理器为 369Mhz ARM，内存 35MB，屏幕显示分辨率为 240×320 像素，内置摄像头采集视频图像分辨率为 198×264 像素。Nokia 6120c 的操作系统为 S60 3rd 9.1 版本，携带的三维图形渲染库位 OpenGL ES 1.1。

4.1 问题分析

ARToolKit 是目前最常用的桌面级 AR 系统的工具包。文献[24]将 ARToolKit 移植到 Symbian 智能手机上。基于此，本文将 ARToolKit 移植到 Nokia 6120c 智能手机上，分别记录跟踪过程中角点检测、标记识别以及摄像机姿态计算三个任务各自的运行时间。在 ARToolKit 中，采用逐帧检测计算法，即对每一捕获的视频帧均全帧检测角点、计算摄像机姿态。ARToolKit 使用结构体 *ARMarkerInfo* 用来存储标记信息，其中 *cf* 表示合法标记的置信度，*area* 表示标记黑色连通区域的像素数。当 $cf = 0.891720$ ，*area* = 9107 pixels 时，连续地记录每 20 帧完成各部分所使用的时间，实验结果如图 4.1 所示。

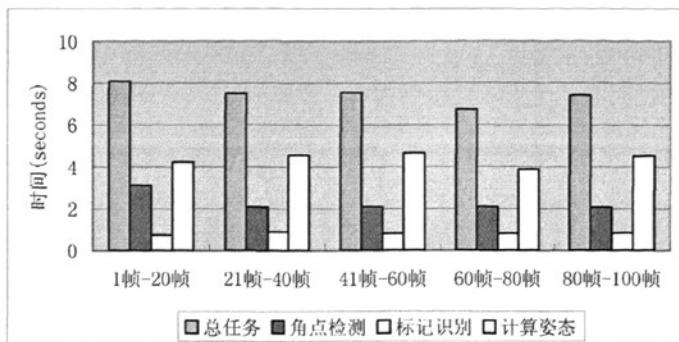


图 4.1 Nokia 6120c 智能手机上使用 ARToolKit 工具包完成跟踪任务的时间

由实验结果可得，角点检测与摄像机姿态计算耗费了大量的系统时间，这对于计算性能较低的手持设备而言是极大的计算负担，影响了系统跟踪注册的实时性。因此，本文将针对标记角点检测与摄像机姿态计算这两方面内容进行研究，在 4.2 节中提出基于区域生长的直接角点检测算法以提高角点检测的速度，在 4.3 节中建立基于窗口的摄像机姿态估计算法，提高跟踪注册的实时性。

4.2 基于区域生长的直接角点检测算法

标记角点位置的检测是标记识别与摄像机姿态计算的首要步骤。在本文中，使用黑白平板标记，标记具有粗的黑色边框，边框内部绘有不同的图案，使得标记具有独立的 ID，如图 3.3 左所示。目前已有的标记检测大多是通过标记的四条直线边相交确定标记的四个角点，或是通过标记边框上最远距离确定角点，本文 3.2 节对这些方法有详尽的描述。但是，这些方法均间接得到角点，计算复杂，计算量也较大。

考虑到标记的边框是较大的黑色连通区域，根据这一特点，本文设计一种基于区域生长的直接标记角点检测算法 (Corners Finding Directly，简称 CFD)，其基本原理为：首先对输入的图像进行阈值分割，利用标记边框具有较大的黑色连通区域的特点，使用区域生长的方法检测出黑色边框，而该区域上具有极值的像素即为标记的角点。

4.2.1 阈值分割方法的选择

本文将 3.2.1.3 小节中讨论的局部动态阈值分割方法移植到 Nokia 6120c 智能手机上，对该方法进行性能分析。

设置窗口大小为 $Image_Width/8$ ($Image_Width$ 为采集到图像的宽), $t = 0.15$, 如图 4.2 显示了在标记面积大约占图像总面积 $1/6$ 和 $1/4$ 两种情况下阈值分割结果:

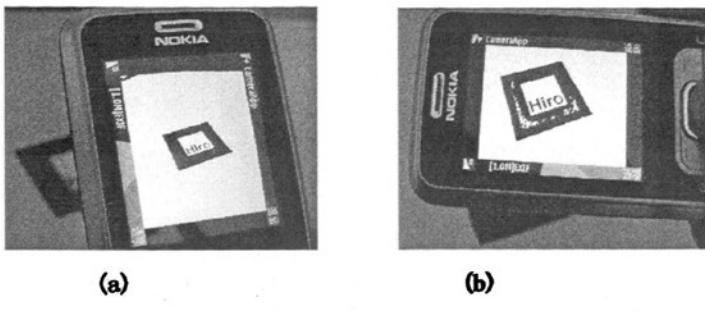


图 4.2 (a) 标记占图像面积 $1/6$ 时, 采用局部动态阈值分割结果;
(b) 标记占图像面积 $1/4$ 时, 采用局部动态阈值分割结果

图 4.2 显示当标记尺寸较大时, 黑色边框部分被分割为背景, 出现了不希望的分割结果。这是因为当标记尺寸较大, 而窗口设置较小, 当前像素小于窗口像素平均值的 $t \times 100\%$ 。因此, 该方法要求随着标记尺寸的不同动态调整窗口尺寸, 并且, 该方法需要两遍扫描源图像, 这对于手持设备上的 AR 系统增加了计算负担, 降低了系统的实时性。

因此, 本文为了提高角点检测的速度, 采用全局静态阈值的方法, 设置全局静态阈值为 100, 将输入图像中像素值与 100 比较, 小于 100 为物体, 大于 100 为背景。在常规情况下该方法是可以满足阈值分割要求的。

4.2.2 快速标记角点检测算法 CFD

按顺时针为标记的四个角点设置编号, 如图 4.3 所示, 分别是 $Corner_1$, $Corner_2$, $Corner_3$, $Corner_4$ 。该黑色连通区域内的所有像素的坐标值分别在横、纵方向上存在极值, 设极值分别为 $MinX$, $MaxX$, $MinY$, $MaxY$, 则 $Corner_1$ 的横坐标为 $MinX$, $Corner_2$ 的纵坐标为 $MaxY$, $Corner_3$ 的横坐标为 $MaxX$, $Corner_4$ 的纵坐标为 $MinY$ 。由此, 可以检测得出四个角点。

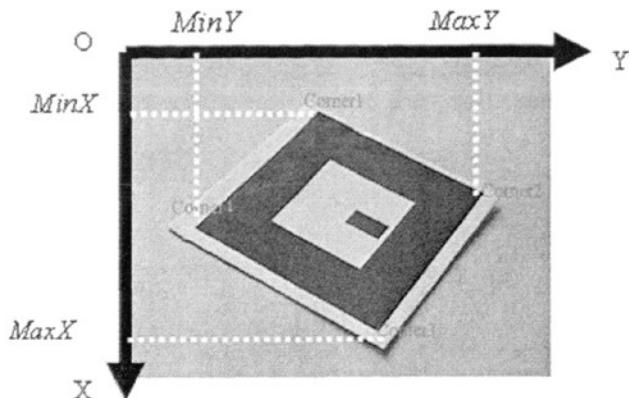


图 4.3 标记的四个角点坐标具有极值特征

具体算法流程为：

Algorithm: CFD

Input: 二值化的帧图像

Output: 四个角点

Begin

```

While(has unlabeled black pixel) do//当存在未处理的黑像素时
    set the unlabeled black pixel as seed = (seedX, seedY); // 设置种子
    Initialize four corners as seed;
    Push the seed into stack;//将种子压入堆栈
    While stack is not null do
        Pop one pixel; //弹出栈顶元素
        For all neighbors of the popped one do //对于栈顶元素的邻居
            If neighbor (NeighborX, NeighborY) is black then//如果邻居为黑
                Push neighbor into stack;\将邻居压入堆栈
                \更新四个角点的坐标
                If NeighborX < MinX then
                    MinX = NeighborX ,Corner1 = (MinX, NeighborY);
                End if
                If NeighborX > MaxX then
                    MaxX = NeighborX ,Corner3 = (MaxX, NeighborY);
                End if

```

```

If NeighborY < MinY then
    MinY = NeighborY, Corner4 = (NeighborX, MinY);
End if
If NeighborY < MinY then
    MaxY = NeighborY, Corner2 = (NeighborX, MaxY);
End if
End if
End For
End While
Set the seed as labeled pixel;           //设置该种子为处理过标志
Check region formed from this seed;   //检测由该种子形成的连通区域
                                      是否为候选区域
End While
End

```

由于可能存在大量的连通区域，需要设置判定规则，保留最可能的候选区域，规则如下：

- ①区域的面积不小于给定面积阈值，排除小区域的干扰。
- ②根据四个角的特点，去掉细长型的区域：当 $Corner_1$ 和 $Corner_2$ 的横坐标（纵坐标）相差小于给定距离阈值，或者是 $Corner_3$ 和 $Corner_4$ 的横坐标（纵坐标）相差小于给定距离阈值，此时的区域必为细长型。

4.2.3 算法性能分析

本文采用 Nokia 6120c 智能手机作为实现平台，摄像头采集到的帧图像尺寸为 264×198 pixels，以检测识别标记为目的，将人们使用手机摄像头的习惯作为阈值设定因素，不考虑标记距离镜头较远、或是处于图像边缘等情况，设置面积阈值为 4500 pixels（大约为帧图像尺寸的 $1/6$ ），可以较好地去除小的连通区域的干扰；设置角点位置之间的距离阈值为 50 pixels，可以去除细长型连通区域的干扰。图 4.4 显示了实验结果，(a) 中去除了标记内部小的黑色连通区域的干扰，(b) 中去除了细长型连通区域的干扰。实验表明，采用本文方法检测角点的时间平均为 29ms/frame。而将 ARToolKit 移植到 Nokia 6120c 智能手机上，检测角点平均时间为 51ms/frame，本文方法可以显著提高角点检测速度。

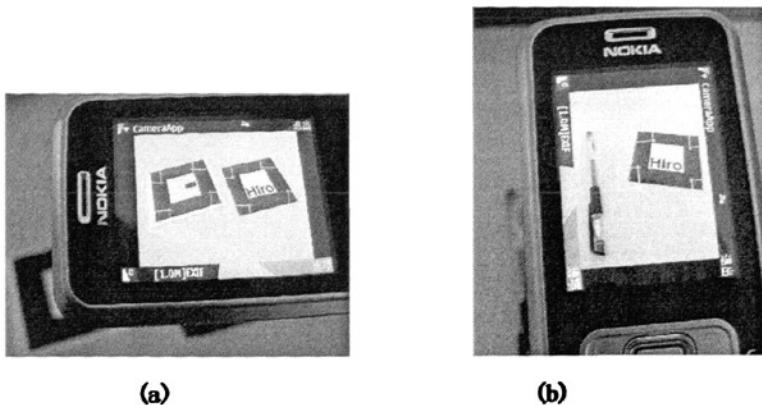


图 4.4 快速角点检测算法实验结果：(a) 去除较小连通区域的干扰；
(b) 去除细长型连通区域的干扰

当然，本算法的前提是图像中的标记的边不能和图像的横纵轴平行，如果平行，则会出现多个极值的问题。但是，该种情况出现可能性极小，即便出现，由于手部的抖动，摄像头姿态很容易会发生改变，并且，该方法是像素级的检测，只要存在不平行的情况，本算法就可以正确的检测出角点。

4.3 基于窗口的摄像机姿态估计算法

4.3.1 问题描述

当用户使用 AR 系统平台时，存在两种摄像机姿态发生改变的情况。其一是用户有意识地晃动手持设备，此时，摄像机姿态变化较大，通常可看作摄像机姿态的改变。其二是用户手部随手持设备无意识地抖动，该种情况主要由以下三方面原因造成：①生理上无意识地抖动；②键盘操作引起的抖动；③行为变化引起的抖动（如行走）。本文仅研究①②两种情况，将①②两种情况定义为手持设备的抖动，主要解决因手持设备抖动影响跟踪注册实时性的问题。为了总结出手持设备抖动的规律，设计以下实验：

在场景中设置标记，用户单手握住手机，使摄像头对准标记，尽量保持手部不动。假设检测到的标记面积大约占整个屏幕的 1/6，在连续的 5 秒钟内记录第 1、2、3、4、5 秒中的标记角点的位置，为简便起见，只列出 $Corner_i$ 的位置。实验对

象共 6 人（3 男，3 女），无任何 AR 系统使用经验，但都有手机使用经验。实验数据如表 4.1 所示。

表 4.1 连续 5 秒内标记的 $Corner_1$ 的坐标

角 点 1 坐 标 用 户	第 11 帧 (第 1 秒)	第 23 帧 (第 2 秒)	第 33 帧 (第 3 秒)	第 47 帧 (第 4 秒)	第 59 帧 (第 5 秒)
用户#1	(61,149)	(60,152)	(63,159)	(63,156)	(57,151)
用户#2	(95,87)	(92,86)	(88,90)	(88,95)	(85,91)
用户#3	(76,83)	(76,80)	(79,86)	(83,85)	(87,76)
用户#4	(57,128)	(55,124)	(56,129)	(49,121)	(49,125)
用户#5	(31,119)	(29,122)	(32,118)	(35,123)	(30,116)
用户#6	(71,75)	(70,75)	(66,79)	(72,73)	(76,81)

由以上实验数据分析可得规律：手持设备随用户手部存在明显的抖动情况，当标记面积大约占整个屏幕的 1/6 时，抖动范围大约为 20×20 Pixels 窗口大小。当手持设备发生抖动时，若仍采用逐帧计算法，则会浪费大量的系统时间，降低跟踪注册的实时性。因此，本文根据提取的规律，提出了基于窗口的摄像机姿态估计算法。其实现机制是：首先判定手持设备是否发生抖动，当确认抖动发生时，无需计算摄像机姿态，使用上一帧摄像机姿态注册虚拟物体；否则重新计算摄像机姿态，更新虚拟物体注册位置。其中，手持设备抖动的判定由连续两帧中同一窗口区域中是否能检测到标记的同一角点加以判断。

4.3.2 算法描述

算法具体实现流程为：当检测出场景中存在合法标记时，记录当前帧中标记角点位置 $Corner_i (i=1,2,3,4)$ 和摄像机的姿态 $TransMat$ 。在下一帧中，仅检测以角点 $Corner_i (i=1,2,3,4)$ 为中心，固定窗口大小区域内是否存在角点，若四个角点窗口中均各自存在角点，则认为手机发生抖动，不计算摄像机姿态，估计摄像机姿态为 $TransMat$ ；否则，重新检测当前整幅帧中是否存在合法标记，计算摄像机姿态。当手机抖动时，为了避免误差累计，设置间隔阈值 $error$ ，初始角点和当前角点位置差值大于 $error$ ，则重新检测当前整幅帧中是否存在合法标记。在窗口范围

内检测是否存在角点时使用的是本章 4.2 小节的直接角点检测算法 CFD 算法。算法流程如图 4.5 所示。

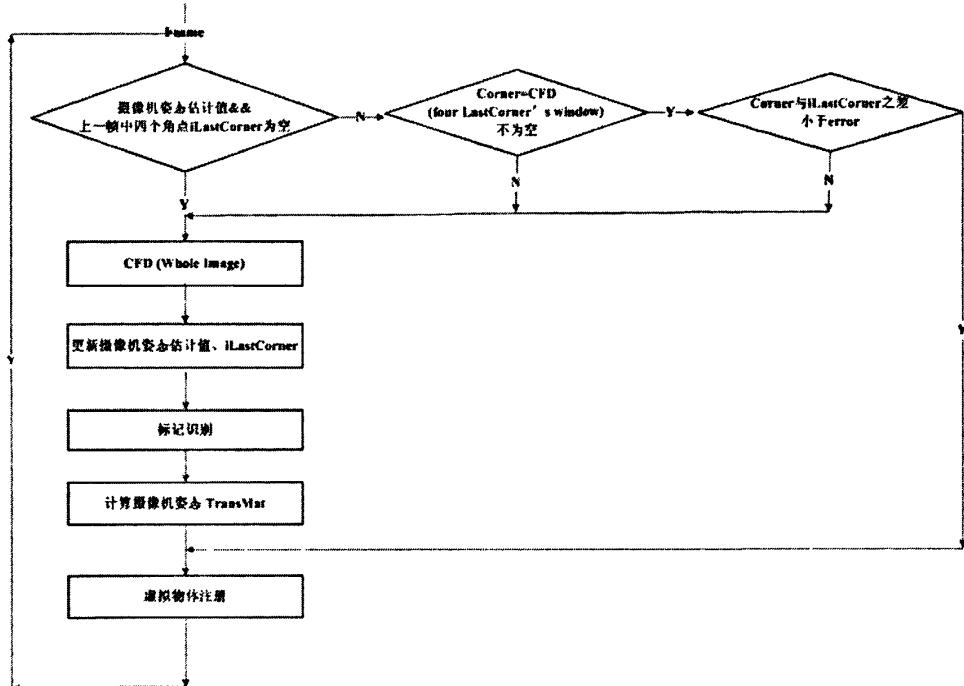


图 4.5 基于窗口估计的摄像机姿态估计算法流程图

具体算法描述如下：

Algorithm: 基于窗口的摄像机姿态估计算法

Input: 二值化的帧图像

Output: 摄像机姿态

Begin

Set TransMat=0, LastCorner_i=0; i=1,2,3,4

Set window size;

Set error;

While (has frame) **do**

 Read one frame;

If (TransMat && LastCorner_i) **then**

```
CFD (LastCorneri's window ((i=1,2,3,4))); //分别在四个窗口内检测角点
If (Corneri (i=1,2,3,4)均存在) then // 如果四个窗口中都各自存在角点, 则认为
    手机发生抖动
    if (Corneri 和 LastCorneri 之差小于 error) then //避免误差累计
        LastCorneri= Corneri; //更新角点
        break;
    End if
    End if
    End if
    CFD (whole frame); //在整帧图像中重新检测角点
    if (pattern_check) then      //为合法标记
        Calculate TransMati;
        TransMat= TransMati, LastCorneri= Corneri; //更新转换矩阵和角点
    End if
End While
End
```

4.3.3 算法性能分析

本文实验环境是：Nokia 6120c 智能手机。要求用户单手持手机，尽量保持手部不动，对准场景中的标记，使标记面积大约占整个屏幕 1/6。本文设置窗口大小为 20×20 pixels，累计误差阈值为 50 pixels。当 $cf = 0.880454$ ， $area = 9485$ pixels (cf 表示合法标记的置信度， $area$ 表示标记黑色边框像素数) 时，连续地记录每 20 帧完成跟踪任务的时间。将 ARToolKit 中逐帧跟踪法与本文提出的基于窗口的摄像机姿态估计算法进行比较，可以发现，由于本文方法首先判断手持设备是否发生抖动，当确认抖动发生时，估计摄像机姿态，极大地提高系统的实时性。实验结果如图 4.6 所示。

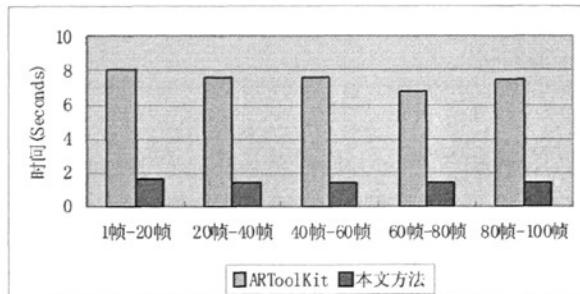


图 4.6 Nokia 6120c 智能手机上分别采用 ARToolKit 与基于窗口的摄像机姿态估计算法进行跟踪的时间性能比较

本文设计的基于窗口的摄像机姿态估计算法中使用固定窗口大小，窗口值与标记占图像面积有关。当标记占图像面积较大时，窗口也应该设置得较大些，同理，当标记占图像面积较小时，窗口也应该设置的较小些。以上实验是在 $area = 9485 \text{ pixels}$ ，即标记面积大约占图像 $1/6$ 大小的情况下得出的。

4.4 小结

本章设计一种基于区域生长的直接标记检测算法，根据标记具有较大的黑色连通区域的特点，通过区域生长的方式，快速地搜索整个连通区域，连通区域上具有极值坐标值的点即为标记的角点。该方法直接检测角点，具有更高的速度优势，对于手持设备是尤为适合的。此外，设计一种基于窗口的摄像机姿态估计的方法，解决了因手持设备随手部发生抖动而引起跟踪实时性降低的问题。通过设置大小合适的窗口，判断手持设备是否发生抖动，当抖动发生时，估计摄像机姿态，否则重新计算摄像机姿态。实验表明当手机抖动时，该方法可显著提高跟踪的实时性。在下一章中，本文将以上设计实现的方法应用于手持设备上的 AR 系统中。

第五章 手持设备上基于视觉标记的增强现实平台

本章设计实现了手持设备上基于视觉标记的增强现实平台，将第四章中研究实现的快速角点检测算法CFD，与基于窗口的摄像机姿态估计算法应用到本系统中完成摄像机姿态跟踪的功能，并且，研究手持设备上三维虚拟图形的渲染，设计实现了三维模型解析器，使得三维模型的复用成为可能，加快了AR系统的原型设计。基于此，实现了Nokia 6120c上的基于AR技术的游戏——打小熊。

5.1 平台结构

本文选用 Symbian OS 智能手机 Nokia 6120c 作为系统硬件平台，处理器为 369Mhz ARM，内存 35MB，屏幕显示分辨率为 240×320 像素，内置摄像头采集视频图像分辨率为 198×264 像素。Nokia 6120c 的操作系统为 S60 3rd 9.1 版本，携带的三维图形渲染库位 OpenGL ES 1.1。

平台的总体结构流程如图 5.1 所示：

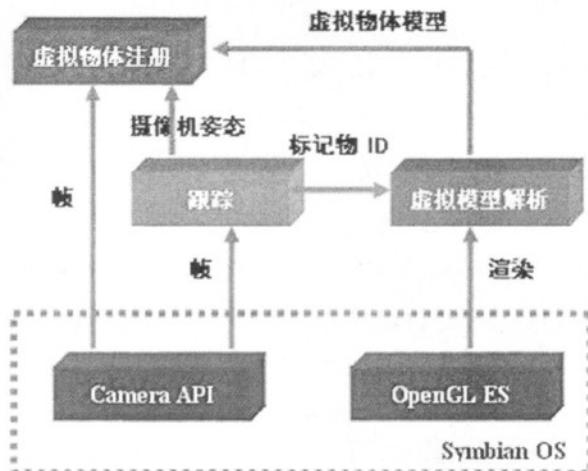


图 5.1 基于 Symbian OS 智能手机上增强现实平台结构数据流程图

整个平台由以下三个功能模块构成：跟踪、虚拟模型解析、虚拟物体注册。各模块的功能描述如下：

①跟踪：跟踪摄像头姿态。对摄像头采集到的视频图像进行处理，检测、识

别标记，获得标记的ID，并且计算摄像机相对于标记的姿态，获得表示摄像头姿态的位置转换矩阵。跟踪共分为三个子功能模块：标记检测、标记识别、摄像机姿态跟踪。基于文献[24]的工作，本文将ARToolKit跟踪库移植到Nokia 6120c手机上，标记识别与摄像机姿态跟踪两个子模块的功能由ARToolKit实现。标记检测子模块的功能由4.2小节设计的快速角点检测算法CFD实现，手持设备抖动影响跟踪注册的实时性的问题由4.3小节设计的基于窗口的摄像机姿态估计实现。跟踪功能模块如图5.2示。

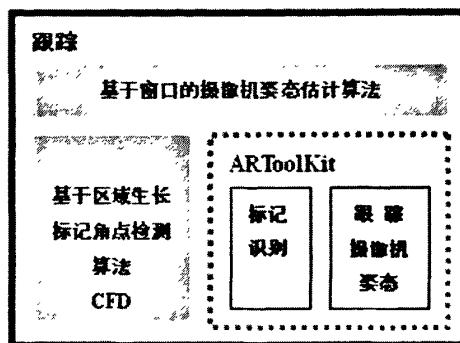


图 5.2 跟踪功能模块

②虚拟模型解析：根据标记 ID，将相对应的模型通过三维模型解析器 Model Parser 转换的模型数据，渲染生成相对应的虚拟物体。

③虚拟物体注册：将采集到的真实场景作为虚拟空间中的纹理背景，并根据转换矩阵实现虚拟物体的注册，最后输出虚实合成视频图像。

5.2 三维模型解析器

5.2.1 COLLADA 简介

COLLADA^[55](COLLAborative Design Activity)是面向交互式 3D 应用程序的基于 XML 的数字资产交换方案，其作用是使 3D 创作应用程序可以自由地交换数字资产而不损失信息。COLLADA 是一个开放的标准，最初用于 3D 软件数据交换，由 SCEA 发起，现在则被许多著名厂家支持如 Autodesk、XSI 等。COLLADA 不仅仅可以用于建模工具之间交换数据之用，也可以作为场景描述语言用于小规模的实时渲染。COLLADA FX 支持使用 OpenGL ES^[42] 着色语言创作和封装着色器，使一流的 3D 创作工具可以有效协作创

建 OpenGL / OpenGL ES 应用程序。

COLLADA 文件后缀是.dae，文件是以 XML 格式组织的，主要有以下几个标签：

<asset></asset>表示文件的基本信息，包括作者、创建时间以及虚拟空间中哪个轴是垂直向上的；

<library_images></library_images>表示场景中采用的纹理图片的信息，包括纹理图片的存储位置；

<library_effects></library_effects>表示场景中的光照和材质，以及纹理参数的信息；

<library_materials></library_materials>对应着<library_effects>中相应的 id；

<library_geometries></library_geometries>表示场景中点的坐标、点对应的向量与纹理坐标，以及各个三角面的点索引、向量索引和纹理索引。

<library_visual_scenes></library_visual_scenes>表示场景由哪些虚拟物体组成，以及虚拟物体的几何变换；

<sense></sense>表示将哪些场景装载进行渲染。

5.2.2 三维模型解析器（Model Parser）的实现

Symbian OS 从 9.1 版本开始自带了 OpenGL ES 三维图形渲染库，支持手机上三维虚拟物体的绘制和显示。为了复用通用建模软件建模得到的三维模型，并且提高 AR 系统开发效率，论文设计实现了三维模型解析器（Model Parser），将通用三维建模软件（如 MAYA, 3DS MAX, Google SketchUp 等）设计建模的三维虚拟物体经 COLLADA 插件导出，得到包含所有可重建模信息的 COLLADA 文件。在 COLLADA 文件中，模型的每个三角面的点坐标、向量坐标以及纹理坐标被设置成了不同的索引，这是不符合 OpenGL ES 绘制渲染要求的。在 OpenGL ES 中对物体进行绘制时只允许使用 VertexArray 数组形式，该数组要求点坐标、向量坐标以及纹理坐标的索引必须一致，因此，需要将 COLLADA 文件中的点数据、向量数据以及纹理数据按照同一索引进行重新排列，使得它们可以符合 OpenGL ES 绘制渲染的要求，三维模型数据解析的流程如下所述：

Algorithm: 三维模型解析

Input: COLLADA 文件 (.dae)

Output: 3-D 模型数据文件

打开.dae 文件；

While (Not EOF) do

 搜索标签<library_images>找到场景中使用的纹理图片；

 搜索标签<library_effects>找到场景中使用的光照、纹理参数以及纹理环境；

 搜索标签<library_geometries>

For 三角形面 i **do**

 抽取出三角形面 i 的点坐标数组 V_i 、向量坐标数组 N_i 以及纹理坐标数组 T_i ；

 抽取出构成三角形面的点索引 $(V_{i0}, N_{i0}, T_{i0}, V_{i1}, N_{i1}, T_{i1}, V_{i2}, N_{i2}, T_{i2})$ ；

End For

 按照点索引重新排列构成三角面的点、向量以及纹理数组，得到 V'_i , N'_i , T'_i ；

End While

5.2.3 实验结果

在 MAYA 8.5 三维建模软件上，分别设计制作了三维模型五月的风、帆船以及水立方，利用软件中 COLLADA Plug-in 插件工具将它们导出为 COLLADA 文件，通过三维模型数据解析器转换这些模型数据，使其符合 OpenGL ES 绘制渲染格式的要求。在 Nokia 6120c 智能手机上注册渲染了这些三维模型，图 5.3 显示了渲染输出的效果。图 5.4 显示了不同渲染条件下手机渲染输出画面的帧率，很明显，随着三维物体三角面片数量的增多，实时渲染的速率降低。

Empty: 无任何 AR 任务；.

Video Render: 仅将背景渲染输出；

Tracking: 采用论文方法检测识别标记并计算摄像机的姿态；

Sailing Boat: 该帆船模型由 310 个三角面组成，共 282 个顶点，使用了一个 64×64 像素大小的纹理图片；

Wind of May: 五月的风模型由 10522 个三角面组成，共 1805 个定点，使用了两个 64×64 像素大小的纹理；

Beijing National Aquatics Center: 水立方模型由 558 个三角面组成，共 457 个顶点，使用了 256×128 像素大小的纹理图片。

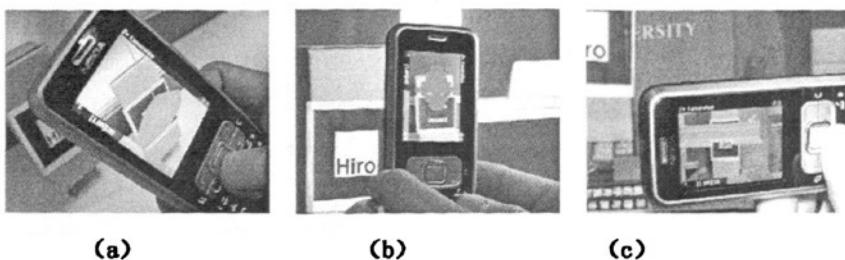


图 5.3 三维虚拟物体注册渲染效果：(a) 为帆船，(b) 为五月的风，(c) 为水立方

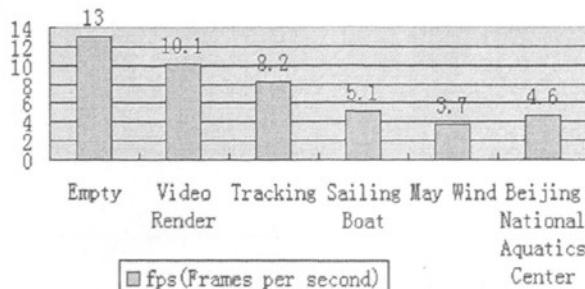


图 5.4 Nokia 6120c 智能手机上注册不同三维物体的刷新帧率

5.3 手持设备上基于 AR 技术的 Bear Hitting 游戏

基于设计实现的手持设备上增强现实系统开发平台，本文设计实现了 Nokia 6120c 上基于 AR 技术的打小熊游戏。游戏时，需要用户在真实场景中设置标记，将手机内置摄像头对准标记，在手机屏幕上标记处将叠加显示三维虚拟小熊，小熊在标记的上下左右四个位置随机出现，用户可通过手机功能键盘控制敲击小熊，图 5.5 为游戏画面截图。

该游戏是基于本文 5.1、5.2 小节设计的平台上开发完成的，其中，图像二值化阈值设为 100，用于估计摄像机姿态的窗口大小设置为 20×20 pixels。三维虚拟小熊的模型在 MAYA 8.5 中设计完成的，考虑到三维模型的精细度对 AR 系统注册的实时性具有较大影响，此游戏中没有为小熊模型设置纹理，小熊模型共有 6964 个三角面片组成。通过 MAYA 8.5 软件中的 COLLADA Plug-in 插件，导出该模型的.dae 文件，由三维模型解析器重新组织模型数据，使其可用于 OpenGL ES 三维图形渲染库使用。

由于本游戏要求用户敲击手机键盘，所以当用户敲击键盘时，手机会发生明

显的抖动，将本文设计实现的方法应用于该游戏具有明显的实时性优势。本游戏仅仅是对本文研究实现的手持设备上跟踪注册技术性能的验证，游戏的功能有限，没有设置积分等功能，这将在以后的工作中实现。

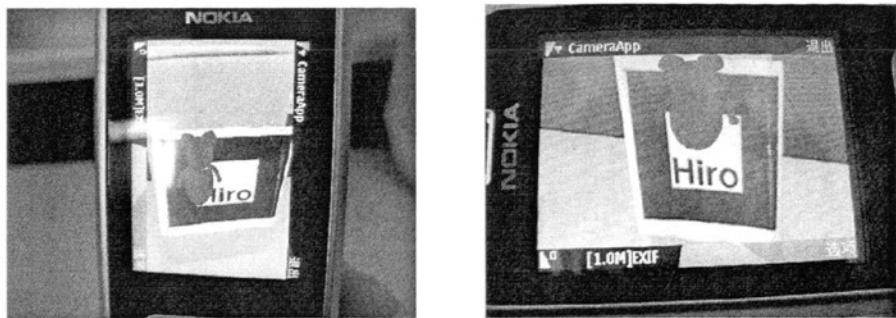


图5.5 Nokia 6120c上基于AR技术的打小熊游戏

结 论

1、全文工作总结

本文研究基于视觉标记的增强现实系统，分析了增强现实系统中的关键技术
和目前常用的开发工具。

- 1) 设计实现了快速的标记角点检测算法，显著提高了角点检测的速度。
- 2) 针对手持设备抖动影响增强现实系统跟踪注册实时性的问题，设计实现了
基于窗口的摄像机姿态估计算法。
- 3) 本文对手持设备上三维图形渲染进行研究，设计实现了三维模型解析器，
使得三维模型的复用成为可能。
- 4) 设计开发了 Nokia 6120c 智能手机上基于增强现实技术的打小熊游戏，游戏
还不具有计分的功能，仅仅对本文设计实现的基于窗口的摄像机姿态估计算法的有效性进行验证。

2、存在的问题及展望

本文虽然取得了一定的研究进展，但还需要在以下方面进行深入研究：

- 1) 针对手持设备移动范围较广，研究对光照具有动态自适应阈值的图像分割
方法，提高 AR 系统的稳定性。
- 2) 本文设计的基于窗口的摄像机姿态估计算法中使用固定窗口大小，窗口值
与标记占图像面积有关。在后期的研究工作中，设计实现窗口自适应的摄像机姿态
估计算法。

增强现实技术是目前新兴的和活跃的一项研究领域，它可以为人们带来新颖的
视觉体验，具有广泛的市场前景和应用范围。尤其是对于游戏产业来说，将增强现
实技术应用于游戏可以极大地丰富游戏的画面内容，增加游戏的娱乐性。

参考文献

- 1 Ronald T. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 1997.6(4):355 - 385
- 2 Milgram P., F. Kishino. A taxonomy of mixed reality visual displays. *EICE Transactions on Information Systems*, No. 12 December 1994:Vol E77-D
- 3 Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.. Recent Advances in Augmented Reality. *Computer Graphics and Applications*, IEEE Volume 21, Issue 6, Nov/Dec 2001 Page(s):34-47
- 4 ARToolKit website: <http://www.hitl.washington.edu/artoolkit/>
- 5 Daniel Wagner. Handheld Augmented Reality: [博士学位论文] Austria: Graz University of Technology, 2007
- 6 Wagner, D., Schmalstieg, D.. First Steps Towards Handheld Augmented Reality. In Proceedings of the 7th International Conference on Wearable Computers (ISWC 2003), USA, 2003:127-135
- 7 PINTARIC, T.. An adaptive thresholding algorithm for the augmented reality toolkit. In IEEE Int. Augmented Reality Toolkit Workshop. 2003:71
- 8 Market Share data from Canalys report. Worldwide smart mobile device market. Canalys Q3 2008.
- 9 Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Anthony Webster. A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. In Proceedings of the 1st IEEE International Symposium on Wearable Computers (ISWC' 97), 1997: 74 - 81
- 10 Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., and Hallaway, D. Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System. *Computers and Graphics*, 23(6), Elsevier Publishers, Dec. 1999:779-785
- 11 Bruce H. Thomas, Benjamin Close, John Donoghue, John Squires, Phillip De Bondi, Michael Morris, and Wayne Piekarski. Arquake: An outdoor/indoor augmented reality first person application. In Proceedings of the Fourth International Symposium on Wearable Computers (ISWC' 00), 2000: 139 - 146
- 12 Adrian David Cheok, Fong Siew Wan, Kok Hwee Goh, Xubo Yang, Wei Liu, and Farzam Farbiz. Human pacman: a sensing-based mobile entertainment system with ubiquitous computing and tangible interaction. In Proceedings of the 2nd workshop on Network and system support for games (NETGAMES' 03), 2003: 106 - 117

- 13 Hirokazu Kato, Mark Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR' 99), 1999: 85 - 94
- 14 Jun Rekimoto and Katashi Nagao. The world through the computer: computer augmented interaction with real world environments. In Proceedings of the 8th annual ACM symposium on User interface and software technology (UIST ' 95), 1995:29 - 36
- 15 Jun Rekimoto. Transvision: A hand-held augmented reality system for collaborative design. In International Conference on Virtual Systems and Multimedia(VSMM' 96), 1996: 85 - 90
- 16 D. Mogilev, K. Kiyokawa, M. Billinghurst, J. Pair. Ar pad: an interface for face-to-face ar collaboration. In CHI ' 02 extended abstracts on Human factors in computing systems, 2002: 654 - 655
- 17 Holger T. Regenbrecht, R. Specht. A mobile passive augmented reality device- mpard. In Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR' 00), 2000: 81 - 84
- 18 Mark Billinghurst, Hirkazu Kato, Ivan Poupyrev. The magicbook - moving seamlessly between reality and virtuality. IEEE Computer Graphics and Applications, 21(3), 2001: 6 - 8
- 19 Christian Geiger, Bernd Kleinnjohann, Christian Reimann, Dirk Stichling. Mobile ar4all. In Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR' 01), 2001: 181 - 182
- 20 Dieter Schmalstieg, Anton Fuhrmann, Gerd Hesina, Zsolt Szalavári, L. Miguel Encarnaçāo, Michael Gervautz, and Werner Purgathofer. The studierstube augmented reality project. Presence: Teleoperators and Virtual Environments, 2002.11(1): 33 - 54
- 21 Daniel Wagner, Istvan Barakonyi. Augmented reality kanji learning. In Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR' 03), 2003: 335 - 336
- 22 Daniel Wagner, Dieter Schmalstieg, Mark Billinghurst. Handheld ar for collaborative edutainment. In Proceedings of 16th International Conference on Artificial Reality and Telexistence (ICAT' 06), volume 4282 of Lecture Notes in Computer Science, 2006: 85 - 96
- 23 Mathias Möhring, Christian Lessig, Oliver Bimber. Video see-through ar on consumer cell-phones. In Proceedings of the Third IEEE and ACM International Symposium on

- Mixed and Augmented Reality (ISMAR' 04), 2004: 252 - 253
- 24 Henrysson, A., Billinghurst, M., Ollila, M.. Face to Face Collaborative AR on Mobile Phones. Proceedings International Symposium on Augmented and Mixed Reality (ISMAR' 05), Austria, 2005: 80-89
- 25 Rohs, M.. Marker-Based Embodied Interaction for Handheld Augmented Reality Games, In Proceedings of the 3rd International Workshop on Pervasive Gaming Applications (PerGames) at PERVASIVE 2006, Ireland, 2006:163-167
- 26 Mozzies Game website:
http://al.siemens.com/innovation/en/publikationen/publications_pof/pof_fall_2004/research_cooperation.htm, 2009-5-20
- 27 Bucolo, S., Billinghurst, M., Sickinger, D.. Mobile maze: a comparison of camera based mobile game human interfaces, Proceedings of the 7th international conference on Human computer interaction with mobile devices & services (MobileHCI' 05) Austria, 2005: 329-330
- 28 Geiger, C., Paelke, V., Reimann, C.. Mobile Entertainment Computing. In Lecture Notes in Computer Science, Vol. 3105 / 2004, Springer Verlag, 2000: 142-147
- 29 Hakkarainen, M., Woodward, C. Symball-Camera Driven Table Tennis for Mobile Phones. Poster at ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE 2005), Spain, 2005: 321-324
- 30 陈靖, 王涌天, 林倞. 增强现实技术在PDA上的应用. 光学技术, 第33卷第一期, 2007. 1:52-55
- 31 陈靖, 王涌天, 施琦, 阎达远. 基于视觉的增强现实运动跟踪算法. 计算机应用研究, 2002. Vol12: 15-21
- 32 施琦, 王涌天, 陈靖. 一种基于视觉的增强现实三维注册算法. 中国图像图形学报, 2002. 7: Vol 7(A):679-683
- 33 哈涌刚, 周雅, 王涌天, 阎达远. 增强现实的头盔显示器的设计. 光学技术, 2000. 7: Vol 26(4):
- 34 徐彤, 王涌天. 增强现实三色立体基准注册技术的研究. 光学技术, 2000. 11: Vol 26(6):535-540
- 35 周雅, 晏磊, 赵虎. 增强现实系统光照模型建立研究. 中国图像图形学报, 2004. 8: Vol 9(8):968-972
- 36 周雅, 徐彤, 阎达远, 王涌天, 哈涌刚. 增强现实中虚拟物体的投影注册算法研究. 北京理工大学学报, 2000. 12: Vol 20(6):738-742
- 37 明德烈, 柳健, 田金文. 仿射变换在增强现实中的应用J. 系统仿真学报, 2001, 13(增刊B): 286-289

- 38 曹达. 手持设备上增强现实技术研究.[硕士学位论文]上海: 上海交通大学, 2007
- 39 姚远. 增强现实应用技术研究: [博士学位论文]杭州: 浙江大学, 2006
- 40 朱森良, 姚远, 蒋云良. 增强现实概述. 中国图像图形学报, Vol. 9, 2004: 767-774
- 41 任鹏, 康波. 增强现实中一种新的基于标志物的跟踪方法. 系统仿真学报, 2009: Vol 21(2):465-468
- 42 OpenGL ES website: <http://www.khronos.org/opengles/>, 2009-5-20
- 43 Michael Rohs. Visual Code Widgets for Marker-Based Interaction. Proceedings of the Fifth International Workshop on Smart Appliances and Wearable Computing, Volume 05, 2005: 506-513
- 44 Fiala, M.. ARTag, a Fiducial Marker System Using Digital Techniques. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR' 05) - Volume 2, 2005: 590-596
- 45 Wagner Daniel, Reitmayr Gerhard, Mulloni Alessandro, Drummond Tom, Schmalstieg Dieter. Pose Tracking from Natural Features on Mobile Phones. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ismar), 2008: 125-134
- 46 Jingtao Wang, Shumin Zhai, John Canny. Camera phone based motion sensing: interaction techniques, applications and performance study. In Proceedings of the 19th annual ACM symposium on User interface software and technology (UIST' 06), 2006: 101 - 110
- 47 Anders Henrysson. Bringing Augmented Reality to Mobile Phones:[博士学位文]. Austria: Linköpings universitet, 2007
- 48 Daniel Wagner, Thomas Pintaric, Florian Ledermann, Dieter Schmalstieg. Towards massively multi-user augmented reality on handheld devices. In Proceedings of Third International Conference on Pervasive Computing (Pervasive 2005), volume 3468 of Lecture Notes in Computer Science, 2005: 208 - 219
- 49 Anders Henrysson, Mark Ollila, Augmented Reality on Smartphones, Augmented Reality Toolkit Workshop, 2003. IEEE International Volume, Issue, 7 Oct. 2003: 27 - 28
- 50 Piekarzki, W., Thomas, B.. Tinmith - An Architecture for Supporting Mobile Augmented Reality Environments. The 2nd International Symposium on Augmented Reality (ISAR), USA, 2001: 177-178
- 51 Daniel Wagner, Dieter Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In Proceedings of 12th Computer Vision Winter Workshop (CVWW' 07), 2007
- 52 Daniel Diggins. ARLib: A C++ Augmented Reality Software Development Kit:[硕士学位论文]英国: Bournemouth University, 2005
- 53 Handheld Augmented Reality website:

参考文献

- http://studierstube.icg.tu-graz.ac.at/handheld_ar/index.php, 2009-5-20
- 54 Bradley, D., Roth, G.. Adaptive Thresholding Using Integral Image. In the Journal of Graphics Tools, Volume 12, Issue 2. 2007: 13-21
- 55 COLLADA website: www.khronos.org/opengles/, 2009-5-20

攻读学位期间的研究成果

- 1 Yi Sui, FengJing Shao, RenCheng Sun, JinLong Wang, A Sequential Pattern Mining Algorithm Based on Improved FP-tree, SNPD 2008, 2008.8.6, index by EI.
- 2 RenCheng Sun, Fengjing Shao, Shujing Li, Yi Sui, “MagicARPhone: An AR-based Game on Phone Offering Personalized Enjoyment”, World Congress on Software Engineering .WCSE 2009 (to appear).
- 3 参与研究青岛市科技发展计划项目“基于增强现实的网络游戏虚拟实境技术”。

致 谢

在论文完成之际，首先感谢我的导师邵峰晶教授在毕业设计期间对我的悉心指导，是她凭借自己在科研领域敏锐的洞察力和先进的研究理念帮我加深了毕业设计的理论深度，挖掘出其中的学术价值。在三年的研究生学习过程中，邵老师一丝不苟的工作态度和严谨求实的治学风范给我留下了深刻的印象，在引领我学习和科研的同时，她也教会我们很多为人治学之道。我的毕业设计正是在邵老师的指导下才得以顺利完成。在此，向邵峰晶教授表示诚挚的谢意，谢谢老师对我的谆谆教导。

其次衷心感谢孙仁诚老师，在他的繁忙教学与博士学习之余，仍然细心指导我的课题研究进展与论文写作，使我能顺利完成硕士阶段的学习。在生活和做人上也给予我们这些师弟师妹们很多帮助和启示，指导和鼓励我们不断向前奋进。

此外，感谢身边所有关心我的同学和朋友，是他们在科研道路上给我帮助，在人生旅途中给我勇气，使我在青岛大学度过了精彩的三年，找到了积极的态度，经历了很多也收获了很多。

最后，感谢我的父母家人，他们对我的关爱和支持是我不断进步的动力。

基于手持设备的增强现实技术研究与应用

作者：隋毅

学位授予单位：青岛大学

本文链接：http://d.g.wanfangdata.com.cn/Thesis_Y1469787.aspx