

Part 1

第一部分

Kinect 基础篇

2011年3月9日,微软宣布 Kinect 自发售以来已经售出了1000多万部,销售额超过15亿美元。与此同时,他们还售出了1000多万套专为 Kinect 设计的游戏。根据吉尼斯世界记录, Kinect 成为了历史上销售速度最快的消费类电子产品,打破了此前由 iPhone 和 iPad 所保持的记录。微软 Kinect 以“你就是主角”(You are the controller)为口号,正在引领着一场人机交互的变革。

进行 Kinect for Windows 开发,首先要配置开发环境。本书将在第2章详细介绍 Kinect 开发对软硬件的要求,并带领读者一步一步进行环境的配置。这部分最后还会结合自然用户界面(NUI),简要介绍 Kinect for Windows SDK 的整体框架。

第 1 章

Kinect简介



Kinect被誉为第三代人机交互的划时代之作。本章将介绍Kinect的基本概念及其发展历程，并简要剖析其结构功能以及体感交互技术的原理。另外，本章最后还会对Kinect for Windows的应用领域进行概览和展望。

1.1 什么是 Kinect

Kinect是Xbox 360外接的3D体感摄影机，如图1-1所示。它利用即时动态捕捉、影像识别、麦克风输入、语音识别等功能，使玩家摆脱了传统游戏手柄的束缚，使用自己的肢体来控制游戏。而任天堂Wii、索尼Play Station Move等同类产品，则需要玩家借助一个或者多个设备才能完成体感互动。



图1-1 Kinect for Xbox 360

作为Xbox 360的外设，Kinect不需要使用任何道具即可完成整个动作的识别和捕捉，它使用了由微软剑桥研究院研发的基于深度图像的人体骨骼追踪算法，而深度图像则是由PrimeSense公司提供的Range Camera技术产生的。此外，Kinect使用一个4-麦克阵列，可以识别3D立体语音。

Kinect的主要识别算法和软件部分都是由微软旗下的游戏工作室提供的。国内外一些所谓的可见光或者红外识别公司，大多是从该工作室获取一些专利权，其产品跟微软的Kinect相比在精度上还有一定的差距。

1.2 Kinect的前世今生

2009年6月1日，Kinect在E3游戏展上首次亮相，它当时的代号是Project Natal。这遵循了微软以城市名作为开发代号的传统，Project Natal是由来自巴西的微软董事Alex Kipman以巴西城市Natal命名的。Natal是拉丁语，英语中有“初生”之意，由此可见，微软公司期望Kinect能够给Xbox 360带来新生。在E3 2009游戏展上，Kinect的骨骼捕捉技术已经可以在30Hz的条件下同时捕捉4个人的48个骨骼动作。

2010年3月25日，微软宣布将在E3 2010期间召开的“初生计划全球首秀”发布会上公布Kinect的发售日期。2010年6月13日晚，这个发布会在格兰中心体育馆举行，会上微软宣布将Project Natal正式命名为Kinect，这融合了kinetic（运动）和connect（沟通）之意。同时微软还宣布，Kinect将于2010年11月4日在北美正式发售。

Kinect在发布仅仅两个月后，就售出了800多万台，吉尼斯世界记录称其为有史以来销售最快的电子消费产品。但是，Kinect并未就此止步。2011年6月，Kinect for Windows SDK beta版发布，这标志着Kinect开始向PC应用领域进军。2011年11月4日，Kinect发布一周年的日子，世界各地的研究人员已经将Kinect应用到了医疗健康、教育、日常生活等各个领域，以探索Kinect技术的无限可能，这就是所谓的“Kinect效应”。此外，Kinect动作捕捉的机器学习技术还荣获了2011年MacRobert Award工程创新大奖。

微软在Kinect for Xbox 360设备的基础上优化了硬件组件，并于2012年2月发布了Kinect for Windows硬件，其固件更适合PC使用。新的Kinect硬件缩短了USB连接线的长度，并支持“近距离模式”（Near Mode）。与此同时，微软还发布了商业授权版的Kinect for Windows SDK 1.0，这意味着开发者可以使用Kinect for Windows硬件，在Windows平台上开发支持手势和语音识别的应用程序，并向实际用户销售这些程序。

对于商业版的Kinect for Windows，微软采用了纯硬件的商业模式，向开发人员和软件商免费提供SDK开发包。这样，所有的使用者都可以将精力投入到研发上，而不必担心支付任何软件的授权费用。

2012年5月，微软发布了Kinect for Windows SDK 1.5版本，该版本支持人脸以及坐姿半身模式的骨骼追踪。借助这些新功能和特性，Kinect应用程序的开发工作变得更加容易和灵活。2012年10月，微软又发布了Kinect for Windows SDK 1.6版本，主要拓展了Kinect for Windows的开发平台，支持在虚拟机、Windows 8系统上进行开发，支持使用最新的Visual Studio 2012开发工具。此外，Kinect for Windows SDK 1.6版本还增加了获取红外图像等功能，并在性能上做了很大提升。Kinect for Windows SDK可能会保持每年一到两次的更新，在功能和性能上也会越来越强大。

1.3 Kinect 体感交互技术原理

初看Kinect,你或许只看到了3个小摄像头,那么Kinect究竟是怎样实现体感交互的呢? Kinect for Windows SDK又有哪些基本功能呢? 本节将揭开Kinect在硬件、软件方面的神秘面纱。

1.3.1 Kinect的结构组成

图1-2给出了Kinect的整体结构。Kinect一共有3个摄像头,中间一个是RGB摄像头,用来获取640×480的彩色图像,每秒钟最多获取30帧图像;两边的是两个深度传感器,左侧的是红外线发射器,右侧的是红外线接收器,用来检测玩家的相对位置。Kinect的两侧是一组四元麦克风阵列,用于声源定位和语音识别;下方还有一个带内置马达的底座,可以调整俯仰角。



图1-2 Kinect硬件结构

1.3.2 Kinect for Windows SDK 简介及功能介绍

2011年6月17日,微软研究院发布的非商业授权版的Kinect for Windows SDK Beta吸引了众多开发者的目光,不过该版本只允许用于研究、测试和实验,不可以发布商业应用。2012年,微软分别在2月、5月和10月接连发布了商业授权版的Kinect for Windows SDK 1.0版本、1.5版本和1.6版本,此举在明确了微软盈利模式的同时,使得开发者可以进行软件开发,并销售开发的应用程序。

Kinect for Windows SDK目前支持Windows 7操作系统和Windows 8操作系统,开发环境使用Visual Studio 2010 Express及以上版本,支持的开发语言包括C++、C#和VB.NET。

Kinect for Windows SDK主要包括以下几个功能。

- ❑ 骨骼追踪:对在Kinect视野范围内移动的一个或两个人进行骨骼追踪,可以追踪到人体上的20个节点。此外, Kinect还支持更精确的人脸识别。
- ❑ 深度摄像头:利用“光编码”技术,通过深度传感器获取到视野内的环境三维位置信息。这种深度数据可以简单地理解为一组利用特殊摄像头获取到的图像,但是其每一个像素

的数据不是普通彩色图片的像素值，而是这个像素的位置距离Kinect传感器的距离。由于这种技术是利用Kinect红外发射器发出的红外线对空间进行编码的，因此无论环境光线如何，测量结果都不会受到干扰。

- 音频处理：与Microsoft Speech的语音识别API集成，使用一组具有消除噪声和回波的四元麦克风阵列，能够捕捉到声源附近有效范围内的各种信息。

1.4 Kinect for Windows 应用领域

目前，国外已经出现了很多使用Kinect开发的精彩应用，比如Kinect试衣镜、Air Presenter演讲软件、Kinect光剑、Kinect街头霸王等。很多创意都可以在MSDN Channel 9的Coding4fun栏目里看到。

在国内，Kinect for Windows SDK Beta发布伊始，微软亚洲研究院便启动了“微软校园菁英计划”之Kinect Pioneer项目，在全国范围内动员微软学生技术俱乐部的同学们集思广益，提交他们基于Kinect的新创意，并向优秀的创意团队提供Kinect设备和技术支持。仅一个多月的开发时间，多个优秀创意团队便成功提交了Kinect创意项目原型，其中包括使用手势进行变脸的3D脸谱虚拟平台、Kinect教学助手、基于Kinect的网上试衣间等。在随后的2012微软精英大挑战Kinect主题上，来自全国30所高校的100余支队伍也积极参与到Kinect for Windows的开发当中，这使得Kinect在中国大学生中得到了全面的推广。

下面简单介绍一下来自西安电子科技大学团队的3D脸谱虚拟平台。此创意将京剧这门传统艺术和新颖的Kinect技术结合到了一起，通过Kinect搭建了一个可以让京剧迷享受虚拟演唱体验的平台：一个提供脸谱、服装和场景的华丽舞台。凭借着Kinect的人体识别和传感技术，用户可以直接跳过繁复的化妆过程，用手势来选择自己喜爱的角色脸谱，生、旦、净、末、丑，一应俱全，选择完毕后，就可以对着屏幕表演喜爱的曲目并录像了，如图1-3所示。当与其他戏迷朋友分享视频时，他们可以欣赏到惟妙惟肖的场景和表演。这样的方式不仅能使老京剧迷们的交流更加便捷，还能让年轻人通过更炫的途径去了解这门生动的国粹。



图1-3 3D脸谱虚拟平台

6 第1章 Kinect简介

2011年12月2日,由微软亚洲研究院举办的Kinect for Windows研讨会在北京召开,吸引了来自众多行业和研究领域的专家学者以及全国各地的大学老师和学生。研讨会就Kinect的体感交互技术及其应用领域进行了交流和讨论,并且展示了国内基于Kinect for Windows SDK在各个应用领域的研发成果。下面选择一些有代表性的项目进行简要介绍。

- 基于Kinect的手语翻译系统。手语翻译系统旨在解决聋哑人与正常人的沟通问题,利用Kinect对肢体动作的实时捕捉,对特定的手语动作进行识别,最终翻译成文字,这样不懂手语的人也可以跟聋哑人正常交流了。目前,该项目的研究已经取得了重大的进展,我们相信结合Kinect强大的体感交互技术,在不远的将来就会看到Kinect手语翻译系统成为聋哑人的得力助手。
- 空中手写。空中手写软件巧妙地利用了Kinect对手部节点的实时追踪,用户只需对着Kinect在空气中比划,便能写出相应的汉字,并输入到计算机中。该项目为未来的手写输入提供了新的思路。
- 虚拟试衣系统。虚拟试衣系统实现了不需用户真正穿上衣服便可看到衣服上身效果的虚拟试衣体验。用户只需站在屏幕前,选择虚拟试衣系统中存储的品牌衣服图片,系统会根据Kinect获取到的骨架数据自适应地穿在用户在屏幕中的影像上,达到轻轻松松试衣的效果。这样不仅使得试衣变得更加便捷、有趣,还能有效减少试衣成本。
- Kinect版水果忍者。Kinect版水果忍者也是利用了Kinect的骨骼追踪技术,将触屏版的水果忍者游戏移植到Kinect的体感控制上。Kinect追踪玩家双手的移动轨迹,玩家只需对着屏幕划动双手便可切下水果,俨然真实版忍者,在娱乐的同时还能锻炼身体。

1.5 小结

Kinect是一种廉价的动作捕捉设备,适用于对动作捕捉精度要求非常严格的领域,这也是其未来发展的方向。另外,现在智能手机和平板电脑的发展非常迅猛,虽然现在看它们和PC会怎样发展,还没有定论,但是受此趋势影响,Kinect以后绝对会趋于小型化,可以断定这是其发展的必经之路。大家都知道,专业领域的产品用量通常不会很大,但是这个领域的技术更新相对较快,随着技术的完善,最终专业领域的技术一定会逐渐应用到消费者领域。互联网的发展历史就是很好的明证,其他许多成功普及的技术也都遵循着这样的发展轨迹。