



北京航空航天大学

第二十四届“冯如杯”

学生学术科技作品竞赛论文

项目编号 _____

院(系)名称 _____

专业名称 _____

作者姓名 _____

学号 _____

指导教师 _____

2014年4月

智能安防系统

摘 要

社会上的各种各样的恐怖事件以及生活中的偷盗抢劫等时有发生，安全的问题就越来越重要，安全问题设计个人、家庭、公司、学校、社会几乎人活动的每一个领域，而现行的大部分安防设备都是用摄像头去实时录像，而安保人员又不可能时时盯着屏幕，所以大部分只是起到取证的目的，kinect 一种人机交互体感设备，除了在游戏领域独领风骚外，我们将它应用于安防，kinect 采集的视频和相关数据用于判断是否有入侵，在夜间也能精准判断，适用于家庭防盗监控、特殊高安全需求地点等地方的安防。智能安防系统的 PTZ 功能提供智能跟踪模式，能在防区内全区监控，避免视觉死角。可通过网络连接，远方实时观看监控画面。

关键词：安防、kinect、智能监控、远程监控

Abstract

All kinds of terrorist events on society and life steal robbery happen from time to time, security problem is more and more important, safety design of individual, family, company, school, social people almost every field, and most of the security devices are made for the current camera to real-time video, and security guards could not always staring at the screen, so most just have the purpose of the forensics, called a feeling of human-computer interaction body equipment, in addition to outside the game categories, we apply it to security, video and related data from the device to determine whether there is invasion, also can accurate judgment in the night, suitable for family security monitoring, special places such as high security requirements site security. Intelligent security systems PTZ function to provide intelligent tracking mode, can monitor, within the protection zones of avoid blind spots. Can be connected through the network, remote real-time monitor to watch screen.

Keyword: security, Kinect, Intelligent monitoring , Remote monitoring

目录

摘 要	i
Abstract	ii
第一章 作品介绍	1
1.1 项目背景	1
1.1.1 智能安防系统	1
1.1.2 国内外发展现状	2
1.1.3 智能监控与传统监控对比	3
1.2 相关技术	4
1.2.1 Kinect	4
1.2.2 基于 Android 的移动平台	5
1.3 设计目标	6
1.4 工作概要	7
第二章 方案设计	8
2.1 总体设计	8
2.1.1 系统工作流程图	8
2.1.2 整体流程图	10
2.2 监控端	10
2.2.1 Kinect	10
2.2.2 云台与单片机	11
2.3 用户终端	12
2.4 服务器端	13
第三章 系统测试	14
3.1 与 ioimage 的对比	14
3.1.1 ioimage 简介	14
3.1.2 ioimage 优势	14
3.1.3 与本作品对比	16
3.2 作品监控效果	16
3.2.1 白天效果	16
3.2.2 黑夜效果	17
3.2.3 存储效果	17
3.3 硬件搭建	18
3.3.1 监控端	18
3.3.2 用户终端	19
第四章 创新点	20
4.1 将 kinect 用于安防系统	20
4.2 PTZ (Pan Tilt Zoom) 功能	20
4.3 人脸识别辅助判断	21
4.4 判断位置特定算法	21
4.5 声源定位技术	22

4.6 触发式录像.....	23
第五章 总结.....	23
第六章 参考文献.....	24

第一章 作品介绍

1.1 项目背景

1.1.1 智能安防系统

随着现代计算机技术、网络技术的飞速发展和人民生活水平的不断提高，人们对生活环境的要求也越来越高。尤其是对家居安全和家居智能化的要求越来越高。高科技的发展不断推动社会信息化的发展，计算机技术、网络通讯技术以及微电子技术的有机结合，使人类迅速迈入智能化世界。伴随着数字化和网络化，智能化成为了一种不可阻挡的历史趋势。家居智能化将成为 21 世纪住宅的必然趋势，也是我国未来智能住宅建筑的重要发展方向。

另外由于人们生活水平的不断提高，人们对于个人的生活环境，比如安全度、舒适度和生活的便利性有了更高更新的要求。居民对个人住房安全性的标准从上个世纪 80 年代起正在逐渐上升。与此同时，社会人口流动性大大增加，社会结构日趋复杂，诸如入室抢窃、偷盗等诸多社会问题的出现，使人们对家庭生命财产和家居场所的安全越来越重视。

目前大体有两种安防系统：一是采用机械器械来进行防盗，一般都防区安装防盗网或防盗窗，但是此种方式并不能有效地减少潜入、偷盗等事故。二是智能化的安防系统，虽然传统的安防系统达到了一定的智能安防目的，但是在安防系统设计中仍然存在着一些问题。首先智能安防系统的控制中心更多地采用一般单片机作为控制中心，然而这种方式存在很大的弊端。因为一般的单片机资源有限，处理能力也很有限，因此很难满足用户需求，而且也不利于系统的维护和升级；其次终端与控制中心的通信更多的是采用有线的连接方式，然而这种连接方式带来了许多弊端和缺点，比如：线路复杂、线路容易腐蚀和不方便维护等；最后控制中心与远程用户之间的通信更多的也是采用网线和电话线的通信方式，并没有将日益成熟的移动通信技术应用其中，给用户带来了许多不便。

针对目前安防系统设计的不足，本系统采用了数据采集模块、数据处理中心和用户终端的整体构架，运用了迅速发展的无线传感网络与移动通讯网络，以固定平台（如计算机）和移动平台（如手机）作为用户端。同时系统采用了用户终

端实时察看和报警控制相结合的方式，既能实时监控，又可以在关键时刻主动报警，确保了防区的安全，又实现了智能化控制。该安防系统具有高灵活性，低功耗，低成本等特点，为智能安防提出了一种可靠的设计方案。

1.1.2 国内外发展现状

随着我国经济的高速发展和居民收入水平的稳步提高，我国的安防产业发展速度明显加快。改革开放以来，由于我国经济发展迅速，居民的生活水平也有了很大的提高，在这个良好的基础之上，使得我国的安防行业迅速崛起。九十年代起，我国的房地产业迅速发展，极大的促进了安防行业的发展，九八年起我国出现了很多产值超过十亿的大型安防企业。截至到 2006 年中国安防企业已达 2 万多家，从业人员 100 万，主营业收入超过了 1000 亿元。

目前我国的安防产业主要有三种类型企业组成：一是生产安防产品的生产型企业；二是设计安装、服务为主的工程企业；三是经销产品的经销商。目前中国安防企业和产业发展的主要特点为：一是该行业得企业规模迅速扩大，而且产业集中度高，目前国内出现了一批高端企业。二是发展速度快，安防行业的营业收入呈直线上涨。三是产品质量稳步提高，基本满足了中国安防市场的一般需求和国际市场中、低端产品的需求。四是企业日渐重视科技创新，根据一项调查显示，安防企业 2005 年技术开发费用占收入的 5%左右，高于全国工业平均水平。五是市场需求旺盛。安防产品的应用领域越来越广泛，随着人们收入的逐步提高，居民对安防产品的需求量不断增加，智能家居，智能建筑迅速兴起势必会刺激安防产品的迅速发展。同时政府提出了“平安社会”、“平安城市”口号，也进一步刺激了我国安防事业的发展。

目前我国的安防产业虽然有了快速的发展，但是仍然存在着一些问题：企业规模仍然偏小、核心技术缺乏、市场竞争秩序问题较为突出。中国的安防企业仍以中小型企业为主，大型企业较少；目前安防产业的核心技术依然在国外，制约了中国安防产业的发展；中国安防产业仍然存在地区垄断现象，假冒仿冒，侵犯知识产权行为仍时有发生，以上这些状况都严重阻碍了我国安防产业的发展。总体来说，我国的安防产业未来的前景还是非常好的。安防产业将会持续快速发展，初步预计未来几年中国安防产品的市场需求量将保持 20%-30%的速度增长，由于目前安防需求的基数已比较大，因此需求总量已相当可观。而且很多企业都不

断加大对研发的资金和人才的投入，技术方面也会日趋成熟。

而在世界范围来看，以色列的智能安防研究走在世界前列。

出于积极对抗战争威胁的需要，以色列军事和安全产业十分发达。自独立以来，以色列不仅经历了大规模战争，而且与恐怖主义的战争几乎从未间断过，多年的高度戒备和安全意识为以色列安防产业的发展提供了强大的推动力，以色列安防企业制造的许多安全系统和解决方案一直处于领先水平。其智能安防领域典型代表 Visonic、EL3000、Dvtel&ioImage 等公司也先后推出了自己的旗舰产品。



图 1 无线防盗报警器

在当今安全风险日益增多的时代，尤其是恐怖活动，有组织犯罪、电脑黑客、盗窃团伙和破坏活动的日益盛行，安防行业在国外迎来了飞速的发展，无论是政府部门，商业机构还是社会个体，安防都成了不可或缺的一部分。国外的安防产业也起步比较早，而且技术先进，资金雄厚，相比国内来看，国外的安防产业更具规模和实力，而且应用更为广泛。

1. 1. 3 智能监控与传统监控对比

智能监控在数据处理速度和效率上与传统监控相比都有相当的优势。对比结果如表 1 所示。

	传统视频监控	智能监控
监控持续性	需要连续长时间盯着屏幕，容易导致安防人员眼睛疲劳，注意力下降。根据研究表明，人眼紧盯屏幕 22 分钟后，注意力就会大幅下降，远低于正常水平。	不需要一直紧盯着屏幕，安防人员只需要在系统警告时进行确认即可。避免了安防人员因为长时间观看屏幕造成疲劳而降低注意力，提高了安防人员实际监控的效果。
监控有效性	人眼观察范围有限，无法同时紧盯多个屏幕；而人脑也无法同时处理多个屏幕的画面。因	所有监视屏幕均由摄像机自动进行智能监控，安防人员只需要对产生警告的摄像机进

	此,安防人员面临大型屏幕墙时无法有效监控。	行确认和处理即可。这样,实际监控范围可提高数十倍,大大提高了监控效率。
监控能力	人眼视觉灵敏度有限,在监控距离遥远、光线不足等监控环境下,人眼无法察觉监控屏幕上的细微变化。	智能视觉摄像机可以识别出人眼无法分辨的细微变化,例如在遥远距离、光线不足、低对比度、环境伪装等情况下的入侵行为和威胁。
监控实际效果	由于种种原因,传统监控的实际效果并不理想,在大多数时候传统监控只能用于事后取证,无法起到预防、预警的作用。	可以侦测并记录出现在监控屏幕内的徘徊行为。在发现入侵者之后可以自动预警,并通过带 PTZ 的摄像机自动锁定目标并进行跟踪,形成视频跟踪接力和摄像机区域联防。这样,安防人员就可以随时掌握入侵者的行踪,并在事发之前进行预警和控制,把损失降到最低。
监控传输通道有效利用率	无论是否有事情发生,都必须将视频传到监控室的屏幕上,无谓的消耗传输线路的带宽,有其不利于大型监控系统。	可设置“告警触发式”传输模式——借助计算机或嵌入式处理器的强大数据处理功能,对视频图像中的海量数据进行高速分析,过滤掉用户不关心的信息,仅仅为监控者提供有关键信息。
监控录像管理效率	无论是否有事情发生,都必须对监控画面进行录像,极大的浪费存储空间。同时,由于录像存在大量的无用信息,导致回放翻看图像时浪费时间。	可设置“告警触发式”传输模式——只对用户指定的入侵行为或威胁事件前后过程进行录像,在同等条件下可以存储更长时间的录像,同时回放翻查录像的效率很高。

1.2 相关技术

1.2.1 Kinect

Microsoft Kinect 是美国微软公司于 2010 年 11 月 4 日推出的 XBOX-360 游戏机体感周边外设的正式名称。微软 Kinect 在正式发布前一直采用“Project Natal”的名称。Natal 这个名字是微软的一项技术代号,在 E3 上正式宣布。Natal 是一个拉丁名称,事实上是巴西东北部的一个城市,在拉丁语中意味“初生”(To be born)。但这只是一个开发代号,如今,该产品正式更名为 Kinect。Kinect 能够即

时的捕捉人体动作，通过影像对动作进行识别;同时也能通过麦克风的输入，完成语音的识别。因此，玩家可以用肢体的动作控制游戏。Kinect 的社群互动功能还能使玩家通过互联网与他人互动。玩家不仅能一起游戏，还能实时分享图片。。

Kinect 由主机和电源适配器两部分组成，主机设备如图 2 所示：



图 2 Kinect

基座和感应器之间有一个电动的马达，在上面的感应器中有一个红外投影仪，两个摄像头，四个麦克风和一个风扇。由图 2-2 从左往右看，最左边 a 是红外光源，b 是 LED 指示灯，c 是彩色摄像头，可以用来收集 RGB 数据，d 是红外摄像头，用来采集物体的数据，即物体距红外摄像头的距离。彩色摄像头最大支持 280 * 960 分辨率成像，红外摄像头最大支持 640*480 成像。电源适配器由数据输出端及电源端构成，其中数据输出端接口为 USB 接口，可直接与电脑主机连接进行数据传输，非常方便;电源规格为:工作电压为 12V,工作电流为 1.08A。表 1 为 Kinect 的主要技术规格参数。

表 1 Kinect 的主要技术规格参数

技术规格	参数
深度距离	最小距离 80cm、最大距离 400cm
视野角度	水平防线 57° 垂直方向 43°
马达旋转角度	上下各 28°
拍摄视频帧速	30fps
深度图像分辨率	QVGA(320 * 240)
彩色图像分辨率	VGA(640 * 480)
声音格式	16KHz 16bit 单声道
声音输入	4 个带有 24bit 的数字模拟信号和消除和处理噪声的麦克风阵列
骨架追踪距离	最小距离 80cm、最大距离 400cm

1.2.2 基于 Android 的移动平台

本系统客户端运行在 Android 平台。

Android 是 Google 于 2007 年 11 月推出的一款开放的嵌入式操作系统。随着 Android SDK 的不断完善,在 2010 年下半年,Android 规模便超越了苹果和塞班,成为市场占有率最高的智能手机系统。

与其他平台相比,Android 最大的优势在于完全开源,任何拥有 Android SDK 的程序员均可开发;且 Android 并不对开发平台进行要求,开发者可以在 Windows、Unix 等平台上进行软件开发,这与 iOS 的仅允许在 Mac 平台开发相比无疑方便许多。此外,Android 包含一些 C/C++库,这些库能被 Android 系统中不同的组件使用。它们通过 Android 应用程序框架为开发者提供服务。

系统移动平台端主要调用了 Android 平台媒体库。这部分库函数主要基于 PacketVideo OpenCORE。该库支持多种常用的音频、视频格式回放和录制,同时支持静态图像文件。编码格式包括 MPEG4, H.264, MP3, AAC, AMR, JPG, PNG 等。

1.3 设计目标

本系统旨在实现对防区的本地平台及移动平台实时监控,并在认定防区闯入者身份不明的情况下,进行警报提示。

系统设计目标如下:

- 1、通过计算机视觉技术,实现对防区内人像的骨骼识别,用以判断是否有人进入;
- 2、对所识别人像进行人脸识别,并与数据库内已有记录进行比对,若在数据库中无查询结果,则系统进行提示;
- 3、系统的服务器端和移动端可以进行通信,为了使用者能够更好地实施观察防区信息,服务器端可以将监控画面以视频流的形式发送至移动平台;
- 4、由于通过移动网络将视频传输到移动平台在现有网络运营机制下成本过高,因此服务器将判断客户端是在局域网内还是通过移动网络进行访问。若客户端通过局域网进行访问,则服务器正常发送视频流;若移动平台未处于局域网内,则服务器将定时发送经智能选择后的监控照片;
- 5、移动平台可以对监控系统进行控制,包括开启与关闭自动追踪功能、手动跟踪功能、报警功能等;
- 6、为了有效地进行犯罪中止,该系统在提供上位机端报警的同时,还

提出了噪声干扰和报警机制，有不明身份者闯入防区时，使用者可以主动打开噪声报警及干扰装置，利用该装备发出强噪声，一方面用来标明防区处于危险状态，另一方面也可借助该噪声对不明身份者造成干扰，使其心理变化，影响其正常行动。

1.4 工作概要

根据所提设计目标，本项目需要软硬件平台同时配合工作。硬件部分主要包括监控端和通信模块平台的搭建；软件部分包括骨骼、人像识别的图像处理，云台驱动、数据传输和图像显示等。主要工作分为以下阶段：

第一阶段：硬件部分选择。通过一系列对比研究，视频采集部分采用 Kinect 平台；云台部分采用单片机控制步进电机；服务器采用 Windows X86 平台（i5 处理器，2.4GHz 主频）。

第二阶段：硬件平台搭建。将服务器与 Kinect 监控端、云台相连，通过串口与云台主控单片机进行通信。

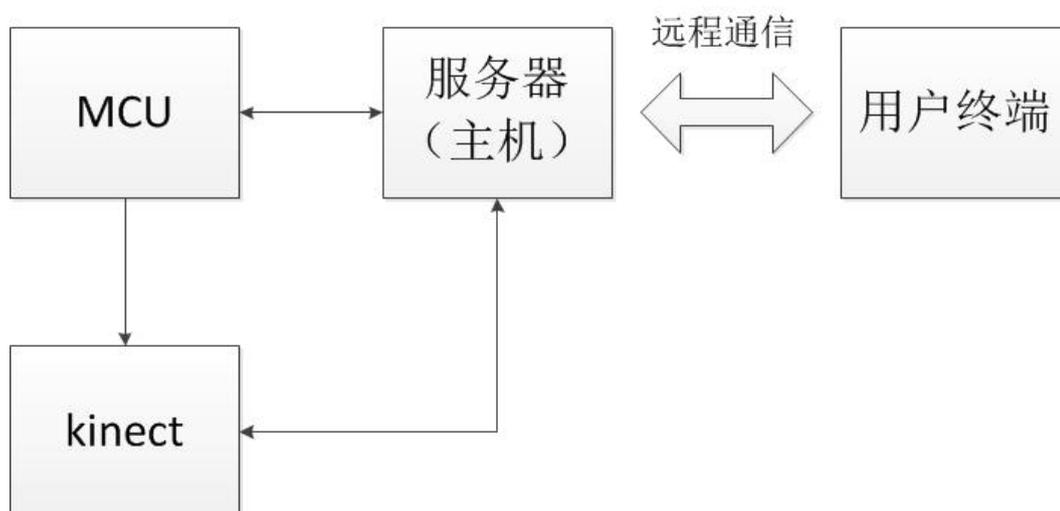
第三阶段：软件部分编写。

- (1) 图像处理部分基于 Visual Studio 2010 可视化编译环境，采用 C# 进行编程；
- (2) 监控云台部分使用 AVR 单片机进行控制，单片机利用 Atmel Studio 环境采用 C 语言进行编程；
- (3) 服务器与移动平台间采用 Socket 通信；
- (4) 服务器端使用 SQL Server 2008 数据库进行人脸数据信息存储；
- (5) 移动平台采用 Java 语言实现 Android 编程，主要利用 Android.io 包中的 ioStream 等结构进行视频流和图片的传输。

第四阶段：系统功能的整体调试。首先进行了 Windows 服务器平台与 Kinect、监控云台、移动平台间的通信测试，确保数据传输工作正常。然后进行了视频流的传输测试，确保局域网内视频数据信息可以流畅的进行传输。最后进行了系统整体功能的测试，包括利用 Kinect 进行视频采集、服务器图像处理、移动平台对监控端的控制等。整个系统工作正常，系统稳定性与处理速度良好。

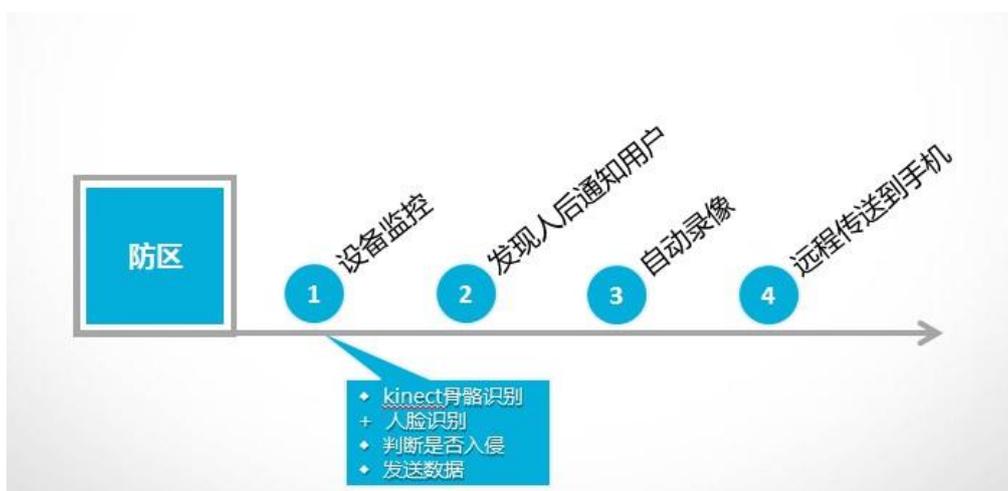
第二章 方案设计

2.1 总体设计

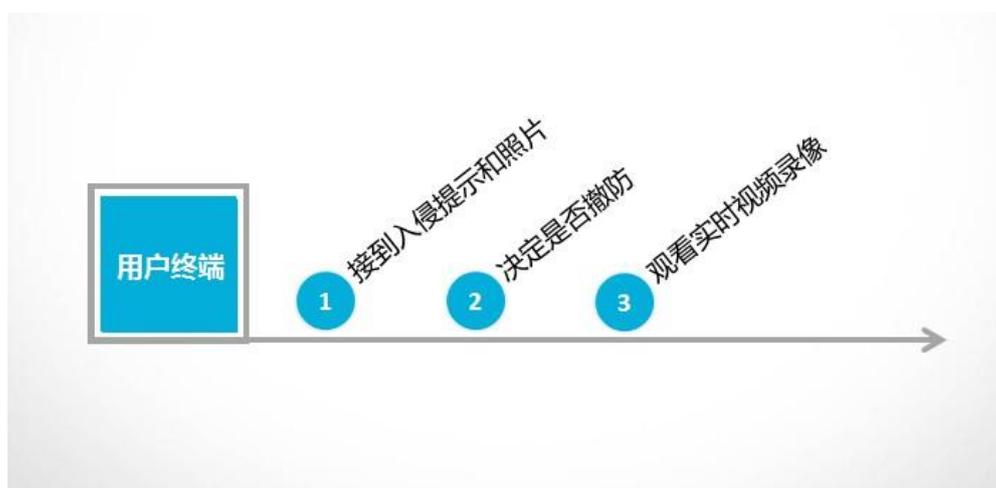


该系统利用 kinect 的骨骼流识别和人脸识别，将二者联合起来判断人的出现及其位置，设置防区后进行监控，方区内出现人后开始录像，用户端的手机即能接到相关提示，人去判断是否撤防，利用 kinect 的俯仰可调和云台的水平可调，监控端可监控并追踪入侵者。

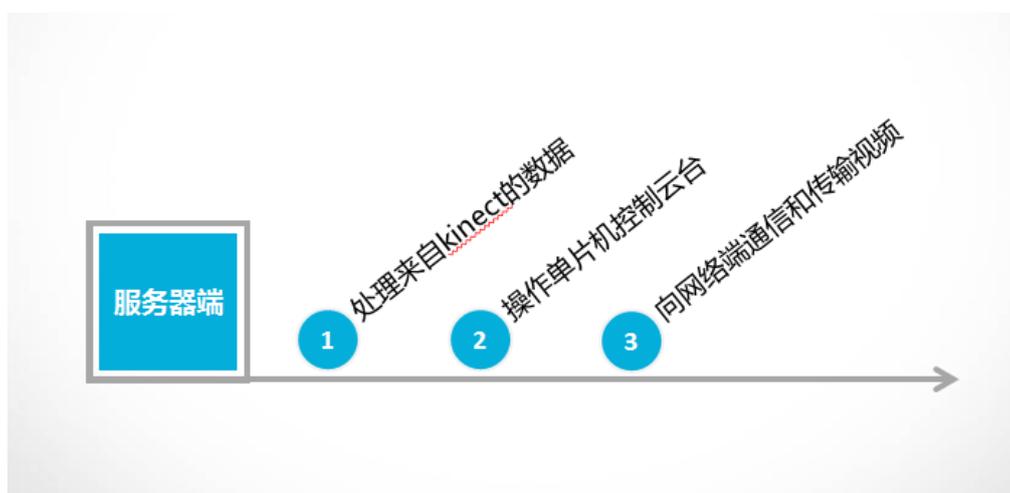
2.1.1 系统工作流程图



防区中用 kinect 去监控防区，当人出现在社区中时，根据 kinect 的编程，利用骨骼流和人脸识别的程序去判断防区人的存在和位置，发现后即向远程的用户发送实时视频和撤防选项，并自动录像和 PTZ 跟踪。

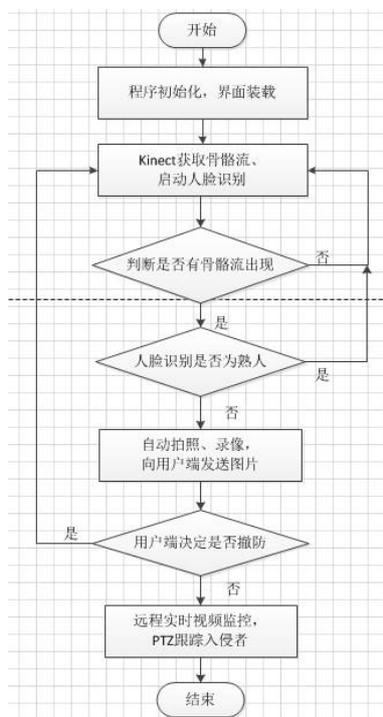


用户终端则在接到入侵提示后让用户决定是否撤防，通过无线传输能实时观看录像，既能在一个局域网下实现该功能，又能在 3G、4G 网络下实现实时视频传输。



服务器端用于处理数据和发送数据，处理来自 kinect 和 mcu 的数据，从 kinect 的骨骼流数据中获取人的肩膀点的位置坐标，用纵坐标控制 kinect 的俯仰，用横坐标控制云台转动的角度，实现 PTZ 跟踪功能，再通过网络端口将实时的视频传送给用户手机端。

2.1.2 整体流程图

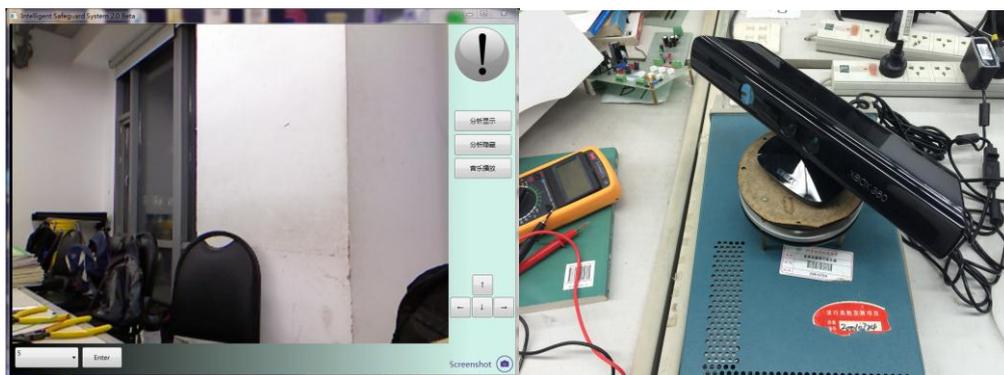


2.2 监控端

该部分包含 kinect 设备、单片机、云台。

Kinect 设备如图：

选择用 C#语言的 WPF 编写程序，用 kinect 和云台作为硬件：



2.2.1 Kinect

Kinect 是微软在 2009 年 6 月 2 日的 E3 大展上。正式公布的 XBOX360 体感

周边外设。Kinect 彻底颠覆了游戏的单一操作。使人机互动的理念更加彻底的展现出来。它是一种 3D 体感摄影机(开发代号“Project Natal”), 同时它导入了即时动态捕捉、影像辨识、麦克风输入、语音辨识、社群互动等功能。

Kinect 使用的是一种光编码 (light coding), 该技术采用连续的照明 (而非脉冲), 也不需要特制的感光芯片, 而只需要普通的 CMOS 感光芯片, 这让方案的成本大大降低。

Light coding, 顾名思义就是用光源照明给需要测量的空间编上码, 说到底还是结构光技术。但与传统的结构光方法不同的是, 他的光源打出去的并不是一副周期性变化的二维的图像编码, 而是一个具有三维纵深的“体编码”。这种光源叫做激光散斑 (laser speckle), 是当激光照射到粗糙物体或穿透毛玻璃后形成的随机衍射斑点。

这些散斑具有高度的随机性, 而且会随着距离的不同变换图案。也就是说空间中任意两处的散斑图案都是不同的。只要在空间中打上这样的结构光, 整个空间就都被做了标记, 把一个物体放进这个空间, 只要看看物体上面的散斑图案, 就可以知道这个物体在什么位置了。

C#中采集骨骼流数据, 在 KinectSensor.SkeletonFrameReady () 中判断是否追踪到骨骼, 即当 skel.TrackingState == SkeletonTrackingState.Tracked 时, 表明已追到骨骼数据, 即在 kinect 的监控防区内有人出现。

但 Kinect 在距离较近时候对人的骨骼识别程度很低, 而人脸识别经测试即使在距离较近的时候仍能有效识别出人脸, 即判断出近处的人, 即将两种数据流同时进行判断, 满足其中一种则表明防区内有人。

一旦发现防区有人, 进行人脸识别, 和数据库中已录入的人脸特征点进行比对, 当相似度超过设置的阈值时则认为入侵者为熟人, 自动撤防, 只向手机端发送一张入侵者的照片和相关操作选项, 如“启动安防”, 如果手机端看到的入侵者并非熟人, 则用户可选择启动安防并报警。如果机器经过人脸识别为陌生人即入侵者, 则报警并传送实时视频并在本地将视频录下以及启动 PTZ 功能。

2.2.2 云台与单片机

在 PTZ 功能中, 我们需要获取人的位置, 并把它反馈给云台和 kinect 设备

的俯仰角，因此我们根据 kinect 的骨骼流数据，从中获取人的肩膀中心的点的坐标，取其 x 与 y ，将 x 的值赋给单片机，单片机通过控制步进电机控制云台，而 y 的值通过控制 kinect 的俯仰角进行调节，达到在防区内的人动，kinect 也跟着动的效果。

通过 C#程序调出 kinect 获取的人肩膀中心点的 x 与 y ，在竖直方向上做如下处理：

(1)、通过人为标定的方式，获取 y 的最大值与最小值，测得大概 y 的最大值和最小值为 0.4 和 -0.4，经实际测量，将角度值除以 0.05 大概为需要变化的角度，当 kinect 正对用户时，其肩膀中心点的 (x, y) 为 $(0, 0)$ ，也可以自己定义一个好的视角，只需要将定义中的 x 与 y 分别加某个坐标值即可。可判断当 y 的坐标大于某个值时，则命令 kinect 的俯仰角发生变化，调整以使 y 变为设定值，如 0。则当 $y=0$ 时表明 kinect 的俯仰角已对准肩膀中点。

(2)、同样通过人为标定的方式获取 x 的最大值与最小值，测得 x 的最大值与最小值为 0.8 与 -0.8，同样将这个值转化为一个需要改变的角度值，将角度值通过 C#的串口通信传递给单片机，经单片机处理后控制云台，让云台产生相应地转动，这个角度值的计算除了需要 x 外还需要 z (kinect 到人的距离)，有了这两个参数才能算出需要改变的角度值。

2.3 用户终端

用户终端可采用手机、电脑等终端设备。我们两套解决方案：

(1)、在安卓手机上运行 APP，通过 JAVA 的编程，完成 socket 通信，当监控端检测到有入侵行为时，手机端弹出消息 让用户确定是否撤防，用户可以在手机端的界面看到监控端采集的图片和实时的视频。

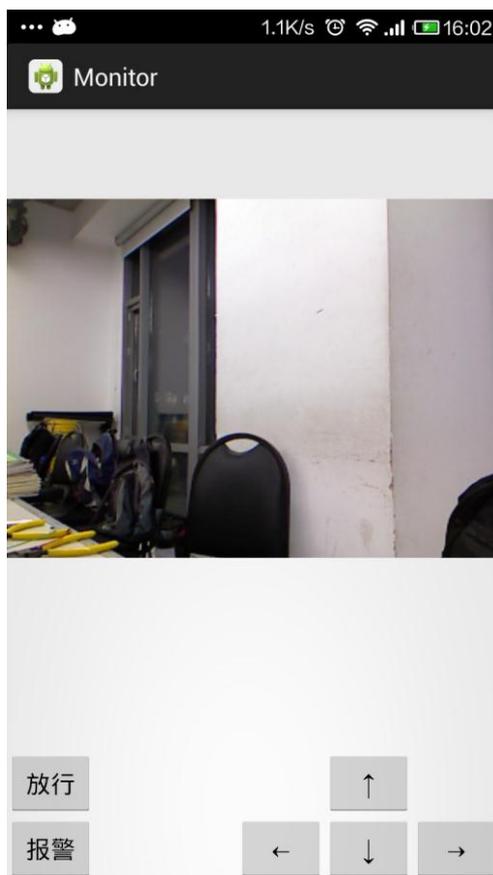


图 3 安卓用户终端

(2)、在远程电脑上运行程序，通过 C#编程，同样为 socket 通信，当监控端检测到入侵行为时，远程电脑端弹出提示，用户确定是否撤防，同时看到传输来的视频和图片，功能大概与手机端相同。

2.4 服务器端

服务器端主要用于采集各种信息、处理各种信息、完成监控端与用户终端的通信以及交互的信息进行反馈等功能，在本作品中，服务器端即指本机电脑，具体的如下：

(1)、对来自 kinect 的数据和来自用户端的数据进行反馈，如 kinect 检测到骨骼流，会把骨骼流传送给服务器端（本地电脑），远程选择“撤防”或者“不撤防”，也会把这个选择信息传递给服务器端，然后服务器端对 kinect、报警铃、录制视频等功能进行控制。

(2)、kinect 把数据发送给服务器端时，经本地电脑判断有入侵现象，则会向远程的用户终端发送提示和传输视频，这部分工作也是由服务器端完成。

第三章 系统测试

3.1 与 ioimage 的对比

3.1.1 ioimage 简介

Ioimage 是一家引领智能视频设备技术的公司，它提供高性能视觉服务器和摄像机，这些产品具有功能强大的内置视频分析技术和独特的设计，并且可组成一整套设备，易于安装和操作。

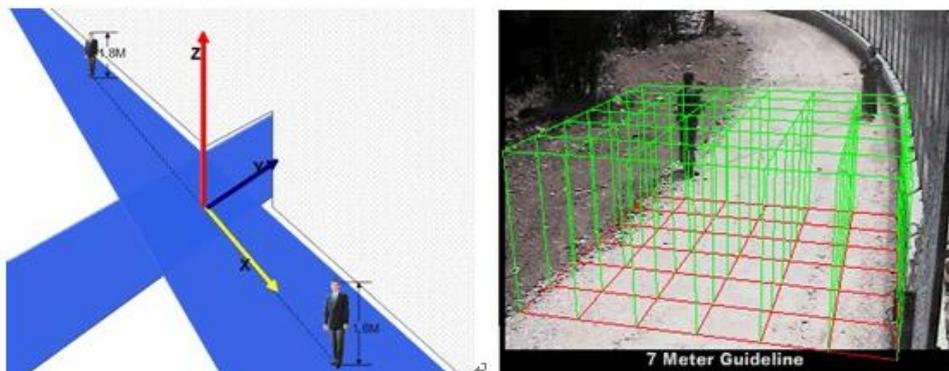
它通过模拟摄像机对采集到的视频做分析，对入侵者、车辆和危险物品进行实时检测、发出报警和跟踪来帮助政府、公众和商业组织将视频监控转换成一种由事件驱动的主动过程。其产品多种多样，从集中、分散和可封装的智能视觉服务器到水平、倾斜和放大（PTZ）智能摄像机，具备用于同步分析可视化的画中画（PiP）跟踪技术以及其他独特的分析功能。

IOI 产品在业界首先提出“景深”概念，并依托设备内的 3D 建模精确的诠释了该概念，且设备内置了 200 多种人物的各种行为特征，能准确的区别人和动物，排除各种干扰物体（如镜头前悬挂遮挡物，飞鸟等）；设备能适应各种地形，山区，丘陵，多弯路地带等；亦能排除各种恶劣天气影响（如：暴雪、暴雨、大风等）。最重要的是设备没有漏报，误报也极低。

3.1.2 ioimage 优势

① 基于 3D 的智能分析技术：

基于 2D 图像进行 3D 建模



② 低漏报、低误报

③ 应用 ioimage 智能视频产品的系统在合理配置的情况下，能够达到不超过

0.1%的漏报几率和不高于1次/3天的误报几率。它可以应用于室内及室外的各种自然环境场合及气候条件下，准确检测入侵。

过滤小动物：



抵抗自然环境因素影响：



④ 特色探测技术

探测极远和极慢目标：



3.1.3 与本作品对比

Ioimage 对远方和大视野的监控很有效，即它是应用于室外监控安防的。而通过实际接触 ioimage 的产品，它在刚进入一个陌生环境中或者镜头转移位置时需要进行标定，即人为地圈定监控头中远方的人，设置大概身高，如 1.75，机器能根据此算出景深，还原 3D 场景。而且它的摄像范围内必须有地面作为参考面才能进行 3d 建模，而且其对近景，如家庭中的卧室、客厅等面积不大的地区，的监控能力有明显的不足，经过测试，其识别率并不高，而我们的作品在近景处出色的表现能力能满足一般情况下的小范围的安防。

3.2 作品监控效果

3.2.1 白天效果

以下为宿舍场景：



以上图表明，刚刚 进入防区即会报警，灵敏度高，而且宿舍环境很复杂（门、窗、衣服、床铺、桌子等），在安防领域已经算是很复杂的监控环境了，但在这种情况下其识别依然很高，而且不需要像 ioimage 那样在新的环境中标定和熟悉

环境的过程。

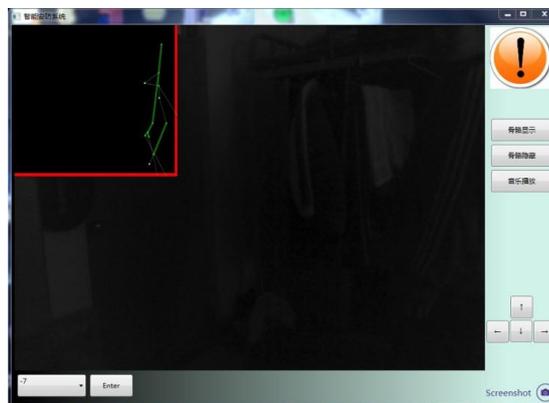
3.2.2 黑夜效果

以下为黑夜弱光场景：

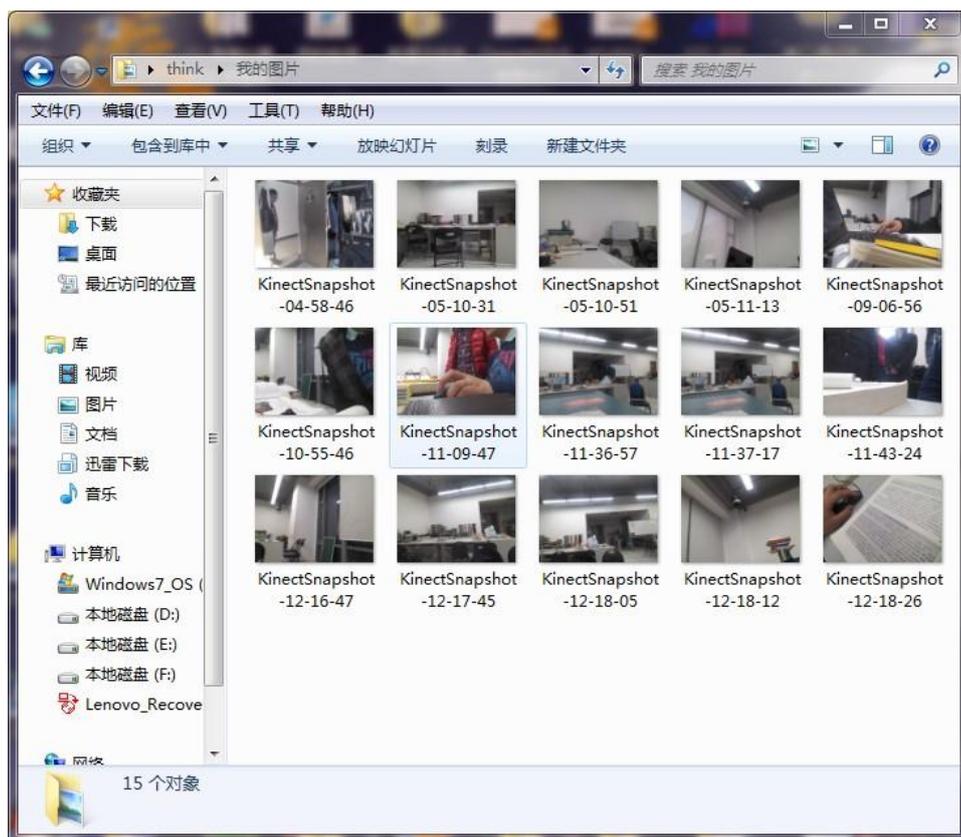


在黑夜弱光场景中，加上宿舍复杂的安防环境，对于一般的安防系统已经不能识别了，而 ioimage 需要复杂的标定和熟悉安防

同过上面两幅图的结果可知，我们的监控设备能比较完美的检测到人，即使在右图中，人眼已经看不到人的位置，而通过左上侧弹窗，我们可以根据绿色小人的位置来判断人的位置，识别率高。



3.2.3 存储效果

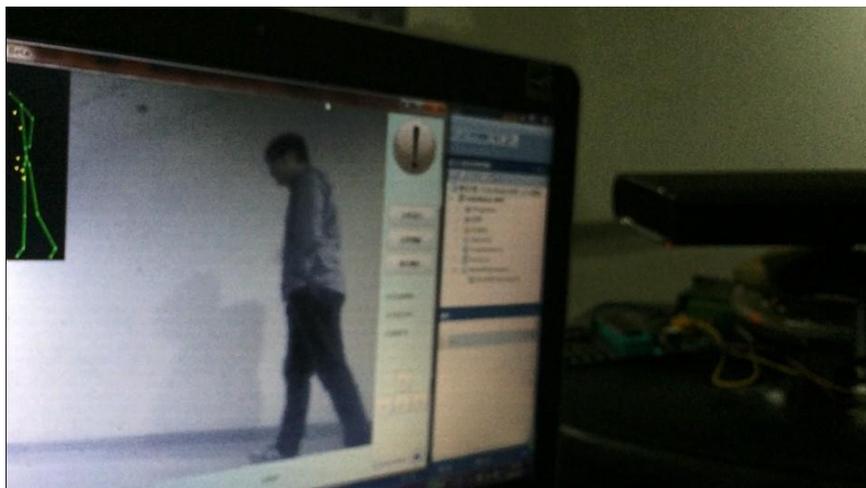


智能安防系统只有在入侵触发后才能录制视频功能，而平时可以通过彩色视频流直接观看摄像头拍的实时视频，节省了存储空间，上图为存储的图片。当遇到入侵行为时，便自动截屏并保存。

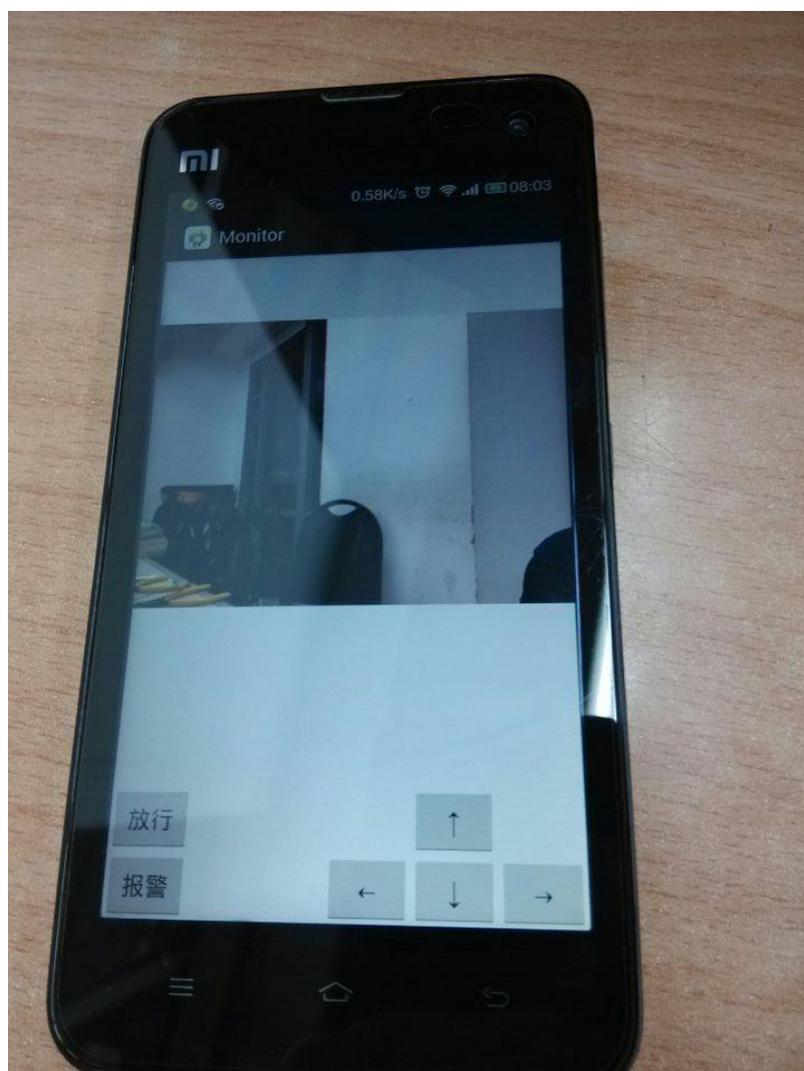
3.3 硬件搭建

3.3.1 监控端

监控端已能实现跟踪人走动，当人进入监控区域时，摄像头会自动调整俯仰角和水平云台的转动，让 kinect 对准人，并跟踪人的行动，当人向左走时摄像头向左偏，人向右走时，摄像头会向右偏，人走进或者蹲下时，俯仰角变小，人走远或者跳起时，俯仰角变大，识别灵敏，而且当两个人同时进入防区时，kinect 设备会采用扫描模式，而当更多人进入时便会选择最大广角去尽可能观察到所有人。



3.3.2 用户终端



第四章 创新点

4.1 将 kinect 用于安防系统

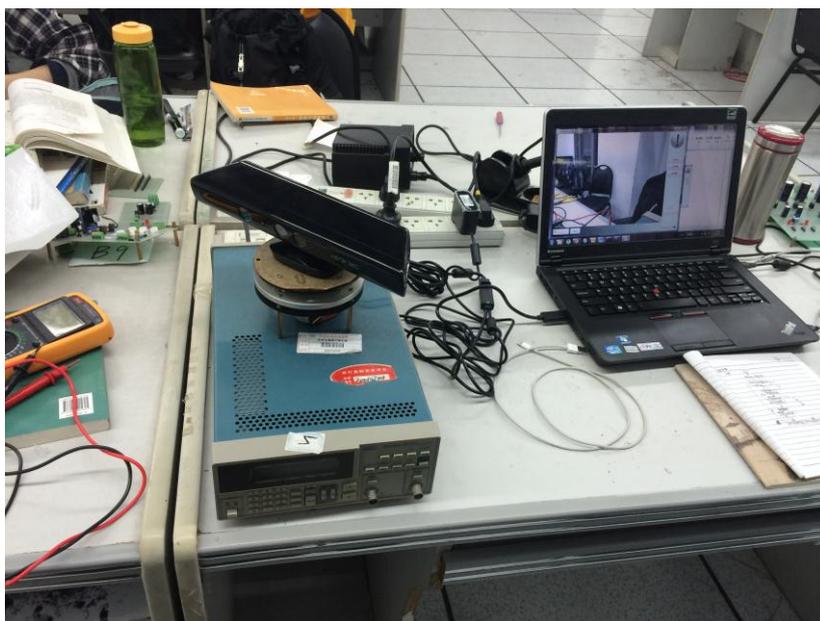
Kinect 的应用场景很多，游戏、交互、操纵电脑等等，但将 kinect 用于安防监控是一个创新点，在网上还没有找到类似的应用场景。

世界范围内最著名的安防公司要数 ioimage 了，上文通过与 ioimage 的对比，可以看到，我们的作品在近场的表现和 ioimage 设备的远场表现几乎相同，而 ioimage 的近场表现却不太理想，而且它需要以地面为参考，拍摄不到地面的话就监控就很不准。

而 kinect 在近场识别人的方面有着很大的优势，它根据反射散斑，和预先写好的算法能精确判断人的位置和行为，经我们测试，效果十分明显，即人一旦走进它的监控视野范围内就能准确判断。

4.2 PTZ (Pan Tilt Zoom) 功能

该系统具有 PTZ 功能，即自动跟踪监控系统，当监控视野中发现人后，将人的坐标反馈给 kinect 和单片机，分别控制 kinect 的俯仰角和水平云台的角度，以使 kinect 的摄像头一直对着人，并且人移动，摄像头也能实时跟踪，达到了扩大视野和跟踪录像的功能。



如上图所示，该图为初级模型，云台也购买了转速更快，更容易控制的舵机 2 自由度云台，在上图中，可实现云台的转动和 kinect 的俯仰变化跟踪人的位置变化。

4.3 人脸识别辅助判断

仅仅根据骨骼识别来判断人的有无有时候会造成误差，比如当人靠近设备时，便无法判断出人的存在，即有近场盲区，为了解决这一问题，我们引入了人脸识别技术，这里所谓的人脸识别指的是获取脸部的特征点，然后对面部进行建模，我们利用的是获取脸部特征点这一过程，采集到人脸特征点后用来辅助判断人的位置，并与数据库里的特征点数据进行对比，发现是熟人后即停止报警，这样，在近场区也能通过判断人脸来识别人的存在，减少了误报率，而且将两种算法同时检测人的存在的算法是独创的。

4.4 判断位置特定算法

为了更好地提取人的位置，写了一套缓冲的优化算法。

人的骨骼流数据或者人脸特征点数据都是一帧一帧出的，而且再加上彩色流数据每次计算机需要处理的数据很多，而且由于人的身体的抖动，其每个关节点的坐标的位置也是不断变化的，如果每一帧都直接传坐标数据的话，不仅处理速度慢，而且还可能出现数据溢出现象，即一次传输的数据太多。

因此，考虑到在这一作品中我们只需要坐标的变化，即 x 和 y 方向的变化量，所以我们在前后两帧中求差值即可，同时降低坐标采集的频率，在 x 和 y 方向的变化上采取导数的形式，当求导后变化很大时，细化这一部分，多采坐标点，同时不断接受反馈的信息进行调整，而当变化不大时便少采点，即加大周期，这样在第一次判别中我们只需要判断其变化量即可，减少了运算量，而且多个线程相互配合，一个数据多次利用的方式也减少了从 kinect 中直接取数据而造成的阻塞。

经过上述一系列优化后，已经明显提高了判断的位置精度和运算速度，跟踪效果也很明显，位置的提取很准确。其中之一的算法如下图所示：

```

//获取坐标的辅助算法
1 private void GetScreenPoint(Joint person, Joint person2, out int w, out int h)
  {
    double w1;
    double h1;
    GetScreenPoint1(person, out w1, out h1);
    double w2;
    double h2;
    GetScreenPoint1(person2, out w2, out h2);
    w = (int)(w1 - w2);
    h = (int)(h1 - h2);
  }
1 private void GetScreenPoint1(Joint person, out double w, out double h)
  {
    w = SystemParameters.PrimaryScreenWidth;
    h = SystemParameters.PrimaryScreenHeight;
    w = w / 2 + person.Position.X * w;
    h = h / 2 - person.Position.Y * h;
  }
  }

```

4.5 声源定位技术

根据仿生学的原理，我们想给监控设备装上“耳朵”，Kinect 的音频系统采用了四元线性麦克风阵列技术。一般而言，麦克风阵列中包含四个相互独立的小型麦克风，每个设备之间相距数厘米，其排列可呈线形，捕捉多声道立体声，通过数字信号处理（DSP）等组件，根据麦克风阵列接听声音的时间差来判断声源方向。



与一般的单麦克风数据相比，Kinect 阵列技术包含有效的噪声消除和回波抑制(Acoustic Echo Cancellation, AEC)算法，同时采用波束成形(Beamforming)技术，通过每个独立设备的响应时间确定音源位置，并尽可能避免环境噪声的影响。

利用声源定位，我们就能知道声音的来源，然后通过单片机控制水平云台让它去“一探究竟”。

4.6 触发式录像

独有的“触发式”录像，只有检测到有人入侵时开启录像功能，避免了普通摄像机一直录像而浪费了大量的存储空间的弊端，只对用户指定的入侵行为或威胁事件前后过程进行录像，在同等条件下可以存储更长时间的录像，同时回放翻查录像的效率很高。

第五章 总结

该系统能提供一套完整的家庭使用解决方案，对入侵者的判断很准确，适用于近场安防，如客厅、卧室等空间。能通过局域网或者 3G 网络与用户保持联系，用户能通过手机或者终端设备来远程监控，当有入侵发生时监控端自动录像、并自动跟踪入侵者，把视频传送到远程客户端，进行实时视频监控。

第六章 参考文献

- [1] 韩旭. 应用 Kinect 的人体行为识别方法研究与系统设计[D]. 山东大学, 2013.
- [2] 基于 Kinect 的人脸图像识别. 戴梅[D]. 湖南大学, 2013.
- [3] 基于 Kinect 的人体目标检测与跟踪. 杨林[D]. 大连海事大学, 2013.
- [4] 陈理. Kinect 深度图像增强算法研究[D]. 湖南大学, 2013.
- [5] 罗鸣. 基于 Kinect 传感器的骨骼定位研究[D]. 武汉科技大学, 2013.
- [6] 林君. 智能家居视频监控系统的设计与实现[D]. 电子科技大学, 2012.
- [7] 李飞霞. 基于 Android 的移动视频监控系统的设计与实现[D]. 电子科技大学, 2013.
- [8] 袁国武. 能视频监控中的运动目标检测和跟踪算法研究[D]. 云南大学, 2013.
- [9] 方建超. 一种新型家居智能安防报警系统的研究[J]. 中国科技信息, 2005(17).
- [10] 武士涛. 基于物联网技术的智能安防系统应用研究[J]. 信息安全与技术, 2012(02).
- [11] 黄凯峰. 智能安防信息化平台的设计与应用[D]. 北京邮电大学, 2012.