

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

工程学导论

Introduction of Engineering



题目：智能物联网管家 Selina 结题报告

姓名学号： 王智航 519021910628

侯烁铭 519021910627

席 望 519021910609

周行远 519021910616

张卓萌 519021910630

朱芷晗 519021910605

指导教师： 焦素娟

智能物联网管家 Selina 结题报告

摘要

在人工智能流行的当下，语音互动、人脸识别等技术逐渐普及，智能管家成为市场中的宠儿。从手机助手（如苹果手机的 SIRI）到智能互动机器人走进家家户户，智能管家不仅满足了人们日常使用的需要，也充实了人们的娱乐生活。本文在该背景下对智能管家的制作进行了研究与尝试。在功能方面，根据人们的日常需要，确定了时间、温湿度、表情等一系列显示功能，并在此基础上添加了简单的运动功能、人脸识别功能与语音识别互动功能，其中最后两项是本项目的特点。在排除多种不可行、不合算的方案后，本项目选择了最优方案，通过 Arduino 单片机控制点阵屏幕显示，树莓派负责人脸识别功能，RecV2 进行语音交互实现预期功能。同时，本项目在机械结构方面进行了详细的设计，通过激光切割的技术得到了所需的亚克力板，扩充了小车空间，并利用小车底盘实现了简单的移动功能。最终拼接出来的样品坚固简洁，功能完好，且对比市场上现有智能互动机器人的单纯娱乐功能，本产品还具有应用到日常生活中的实际功能。

关键词：物联网，语音互动，人脸检测，传感器，闹钟

SMART IOT BUTLER SELINA

CLOSING REPORT

ABSTRACT

In the current era of artificial intelligence, technologies such as voice interaction and face recognition are gradually becoming popular, and smart housekeepers have become darlings in the market. From mobile phone assistants (such as the SIRI of Apple mobile phones) to smart interactive robots coming into every household, smart housekeepers not only meet the needs of daily use, but also enrich people's entertainment life. This article researches and tries on the production of smart housekeeper in this background. In terms of functions, according to people's daily needs, a series of display functions such as time, temperature and humidity, and expressions have been determined. On this basis, simple sports functions, face recognition functions, and voice recognition interaction functions have been added, of which the last two are Is the characteristic of this project. After ruling out many infeasible and uneconomical solutions, the project chose the best solution. The dot matrix screen display was controlled by an Arduino microcontroller, the Raspberry Pi was responsible for face recognition, and RecV2 performed voice interaction to achieve the expected function. At the same time, the project has carried out a detailed design in terms of mechanical structure. The laser cutting technology has been used to obtain the required acrylic plates, which has expanded the space of the trolley, and realized simple movement functions by using the chassis of the trolley. The final spliced sample is sturdy and simple, and the function is intact. Compared with the simple entertainment function of the existing intelligent interactive robots on the market, this product also has practical functions applied to daily life.

Key words: IOT, voice interaction, face detection, sensor, alarm clock

目 录

第一章 绪论.....	1
1.1 智能产品开发现状.....	1
1.1.1 Cozmo	1
1.1.2 Vector	2
1.2 开发平台现状.....	2
1.2.1 OpenCV	2
1.2.2 esp8266	3
第二章 方案设计.....	4
2.1 项目目标.....	4
2.2 系统原理设计.....	4
2.3 方案设计与分析.....	5
2.3.1 机械结构方案.....	5
2.3.2 车轮方案.....	5
2.3.3 视觉方案.....	5
2.4 方案选择.....	6
第三章 详细设计.....	7
3.1 各部分详细设计.....	7
3.1.1 底盘系统.....	7
3.1.2 动力系统.....	7
3.1.3 电子控制系统.....	7
3.1.4 高级传感器系统.....	8
3.2 系统机械结构设计.....	10
第四章 样机组装、测试与评估.....	12
4.1 系统组装.....	12
4.2 系统测试与评估.....	13
第五章 项目总结.....	14
5.1 总结.....	14
5.1.1 表情交互单一	14
5.1.2 机械结构草率.....	14
5.2 展望.....	14
5.2.1 削减成本.....	14
5.2.2 机电一体式设计	14
参考文献.....	15
谢辞	16

第一章 绪论

20世纪四五十年代，第三次科技革命的出现，将我们拉入了“信息时代”之中。我们的生活发生了天翻地覆的变化，互联网的出现与普及极大地造福了社会。

进入了21世纪，科技的动车驶上了高速轨道。革命仍在继续，我们正随着物联网的普及应用而发生点点滴滴的改变——万物互联的理论正在重塑人们的生活方式。打开手机，我们希望可以远程看到家中的情况。我们希望机器可以变得越来越智能，帮助我们免除劳动。许许多多前人看来似乎是魔法的概念，如今也越来越深地渗入了我们的生活：语音交互、人脸识别、大数据、人工智能……我们的项目结合时代发展特点，利用一些新兴的技术，包括物联网、语音、人脸检测等，希望能够设计出具有实际用途的物联网管家。

1.1 智能产品开发现状

1.1.1 Cozmo



图 1-1 Cozmo 宣传图

图 1-1 为由 Anki 研发的智能玩具机器人 Cozmo。Cozmo 个头不大，身高只有 6 厘米左右，外形有点像迷你版的推土机。但它具备人类一样的性格，是一只能和人类进行社交的智能玩具机器人。它的面部是一块有机发光显示器（OLED），眼睛则是随心情变化的方块。因为外形相似，Cozmo 被认为是皮克斯动漫《机器人总动员》中瓦力（WALL-E）的现实版。

每只 Cozmo 由 300 个不同的部分组成，具有复杂的人工智能系统，并融合了机械工程、人工智能、计算机视觉、动画、音响、游戏设计等多个方面。当 Cozmo 充电时，它处于休眠状态，会传来轻微的打鼾声。Cozmo 会自动熟悉所处的环境，一段时间后，用户会发现自己的 Cozmo 越来越聪明。同时，官方发布了 Cozmo 的 SDK，意味着我们可以在买到 Cozmo 后对其进行二次开发。这些都增强了 Cozmo 的可玩性。

1.1.2 Vector



图 1-2 Vector 宣传图

Vector 是一个深灰色的 $2.8 \times 2.4 \times 1.8$ 英寸（HWD）机器人，带有油箱踏板和一个圆形的旋转头。它的头部有一个光滑的黑色塑料正面，覆盖了作为面部的彩色 OLED 屏幕，默认显示大而富有表现力的 Wall-E 风格眼睛，并在被问及时提供更多信息。它具有相同的叉车式臂，可以在前面伸出，用于拾取附带的 Vector Cube 玩具，就像 Cozmo 一样。

除了 OLED 眼睛，Vector 的脸部还有一个高清摄像头，可以直观地看到周围环境。这款相机背面配有四个麦克风，可用于听力，还有一组触摸传感器和加速度计，让它在周围滚动时“感觉”。它甚至还有一个声音，头顶有一个内置扬声器，还有一个文字转语音引擎可以让它说话。

Vector 在一个附带的基站上充电，其两个触点与机器人底部相遇，位于其油箱踏板之间。它可以在电池电量不足时自动寻找其基站进行充电，或者被告知“回家”并且它将重新接通联系人。基站的连接电缆终止于 USB 连接器。

虽然 Vector 外观可能与 Cozmo 看上去很相似，但 Vector 扮演了一个与 Cozmo 非常不同的角色，他是作为一个人永远的伴侣角色，而不是偶尔为了转移孩子们注意力的玩具。他需要变得更智能，识别他所处的环境，并从与他共享的人类家庭中获取线索。

1.2 开发平台现状

1.2.1 OpenCV

OpenCV 于 1999 年由 Intel 建立，如今由 Willow Garage 提供支持。OpenCV 是一个基于 BSD 许可（开源）发行的跨平台计算机视觉库，可以运行在 Linux、Windows 和 Mac OS 操作系统上。它轻量级而且高效——由一系列 C 函数和少量 C++ 类构成，同时提供了 Python、Ruby、MATLAB 等语言的接口，实现了图像处理 and 计算机视觉方面的很多通用算法。

OpenCV 拥有包括 500 多个 C 函数的跨平台的中、高层 API。它不依赖于其它的外部库——尽管也可以使用某些外部库。

OpenCV 为 Intel® Integrated Performance Primitives (IPP) 提供了透明接口。这意味着如果有为特定处理器优化的 IPP 库，OpenCV 将在运行时自动加载这些库。

OpenCV 致力于真实世界的实时应用，通过优化的 C 代码的编写对其执行速度带来了可观的提升，并且可以通过购买 Intel 的 IPP 高性能多媒体函数库（Integrated Performance Primitives）得到更快的处理速度。

1.2.2 ESP8266



图 1-3 ESP8266 实物图

ESP8266 使用了 3.3V 的直流电源，体积小，功耗低，支持透传，丢包现象不严重，而且价格超低。官方提供的 rom 主要有两个，一个是支持 at 命令修改参数的 at 系列 rom，使用此 rom 时，可以使用 at 命令来设置芯片的大部分参数，同时也可将芯片设置为透传模式，这样 ESP8266 就相当于在互联网和 UART 之间架起了一座桥梁。；另一个是物联网的 rom，此 rom 可以通过命令来控制 ESP 的部分 GPIO，而且 ESP8266 也可以采集一些温湿度传感器的数据，然后发送到互联网上。

第二章 方案设计

2.1 项目目标

- (1) 实现基本闹钟功能，在需要时进行时间和温湿度的播报
- (2) 实现简单运动和旋转
- (3) 根据语音等信息显示动态表情
- (4) 进行人脸位置检测，通过旋转实时面向人脸
- (5) 进行特定语音识别，实现简单互动
- (6) 实现物联网功能
- (7) 每次充电后，保证至少三天的续航时间

2.2 系统原理设计

整个系统包括底盘系统、动力系统、电子控制系统以及高级传感器系统。该系统的原理流程图如图 2-1 所示。

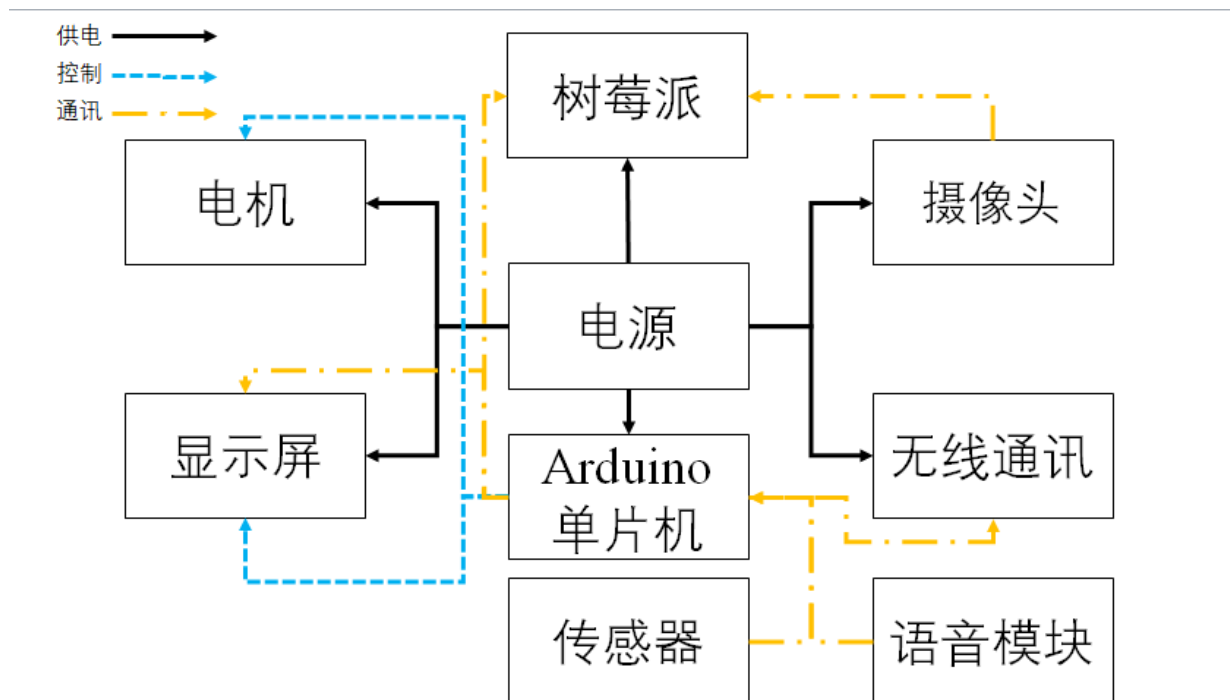


图 2-1 系统原理流程图

整个系统以电源为核心，给底盘系统、电子控制系统以及高级传感器系统进行持续供电。Arduino 单片机作为整个系统的重要组成部分，与显示屏、树莓派、语音模块、传感器进行通讯，实现人脸检测、语音互动、表情显示、温湿度检测显示等功能的结合；同时，Arduino 单片机作为控制中心，通过对通讯信息进行整合，控制显示屏显示表情、温湿度以及控制电机的运作。

2.3 方案设计与分析

2.3.1 机械结构方案

方案一 对闹钟进行改装。通过在市场上直接购买闹钟，对其进行一定的功能改装，加入传感器等元件，实现预定功能。

方案二 对智能小车进行改装。购买常见的智能小车，并通过在小车上布置电路与元件实现预定功能。

机械结构方案分析见表 2-1。

表 2-1 机械结构方案分析

方案名称	优点	缺点
改装闹钟	内置时间显示的功能,且外形较为美观	闹钟体积小,内部结构复杂,改装难度较大
改装小车	能够很好实现运动与旋转功能,且可改造空间大	体积太大,不够美观

2.3.2 车轮方案

为了实现旋转和运动功能,选择合适的车轮是必要的。

方案一 普通车轮。使用常见智能小车的普通车轮,实现物联网管家的简单运动和旋转。

方案二 全向轮。全向轮能够在许多不同的方向移动,结构简单,适用范围广,且能在较差的路况上运动,可以很好地满足物联网管家的运动和旋转的功能。

方案三 麦克纳姆轮。可以实现前行、横移、斜行、旋转等一系列组合运动,适用于在有限的空间内进行使用。

车轮方案分析见表 2-2。

表 2-2 车轮方案分析

方案名称	优点	缺点
普通车轮	便宜耐用	只能实现简单的运动和旋转,且运动和旋转的范围有限
全向轮	可以实现灵活的运动	造价高,且需要对小车进行改造
麦克纳姆轮	可以实现灵活的组合运动,并且车轮较为美观	造价高,且需要对小车进行改造

2.3.3 视觉方案

方案一 树莓派+OpenCV。通过树莓派提供 Linux 系统环境,运行 OpenCV 对摄像头得到的图像进行高效处理,达到人脸检测的功能。

方案二 OpenMV。OpenMV 是基于 Python 的嵌入式机器视觉模块,易拓展,开发环境友好,且内置人脸检测等算法,能够较快实现人脸检测的功能。

方案三 maix go。功能与 OpenMV 几乎相同,能够较快实现人脸检测的功能。

视觉方案分析见表 2-3。

表 2-3 视觉方案分析

方案名称	优点	缺点
树莓派+OpenCV	成本低，且 OpenCV 具有较多的功能	需要进行 OpenCV 的学习，且需要花费较长时间编写代码
OpenMV	易拓展，内置功能多	处理能力有限，且成本较高
maix go	内置功能多，且成本较 OpenMV 低	处理能力有限

2.4 方案选择

综合以上因素，最终选择了以下方案：

通过改造智能小车作为系统的机械框架，用于固定各个元件，同时结合麦克纳姆，实现较为灵活的运动和旋转功能。小车的各个功能通过传感器以及结合 OpenCV 的树莓派实现，并最终通过单片机进行控制。

第三章 详细设计

3.1 各部分详细设计

3.1.1 底盘系统

作为整套系统的基础，底盘的质量决定了整套系统的动作执行能力。

我们选择了 TT 电机作为系统的动力源。6V 供电时，其空载电流小于 0.2A，不超过系统可提供的电流上限。电机空载转速在 200rpm 左右，经过 1:48 的减速箱减速后，能够满足转速与扭矩要求。

为了增强物联网管家的交互性，我们选用了麦克纳姆轮，如图 3-1 所示，在较低设计与制造的成本下使得全向移动成为可能。整车质量小于 2kg，可以不必考虑麦克纳姆轮的磨损严重的缺点。



图 3-1 麦克纳姆轮

3.1.2 动力系统

动力系统采用 5000mAh 充电宝，可持续以 5V/2A 的速率放电，提供的最大功率大于整体 8.5W 的峰值能耗。5V 供电电压的兼容性较高，能满足不同模块的供电需求。电量可以在不进行功耗管理的情况下支持 72h 以上的待机，续航时间可以得到保障。

3.1.3 电子控制系统

电子控制系统采用多核心的拓扑结构，传感器驱动、电机驱动、物联网功能分别由不同的主控系统完成，这种结构可以提高系统的鲁棒性；各主控系统之间采用波特率 9600/115200 的 UART 协议通信，以便进行信息的交流与整合。UART 协议在可以快速交互信息的同时，兼顾了硬件层的简单性与通信的可靠性。

温湿度—时间—彩屏模块以 Arduino Mega2560 为核心，处理速度 16MHz，可以满足时钟模块 DS3231 的 IIC 协议与温湿度传感器 DHT22 的单线协议的时序要求。单个 I/O 口支持高达 40mA 电流，完全满足需求。屏幕的全彩 LED 灯珠采用 ws2812b_5050 灯珠，亮度

高，发热小，控制简单，显色稳定。整屏电流在 0.2A 左右，不会对动力系统造成负担。

电机驱动系统以 Arduino UNO 为主控，配合电机拓展板实现直流减速电机的正、反转。电机拓展板基于直流电机驱动芯片 L298N，可提供 2-4A 的电流，兼具正、反转驱动的实现方式，操作简单，可靠性极高。驱动信号以 PWM 的形式输入电机拓展板，从而实现电机的调速，是 TT 减速电机的一种经典驱动方案。

以上电子控制系统的电路图如图 3-2 所示。

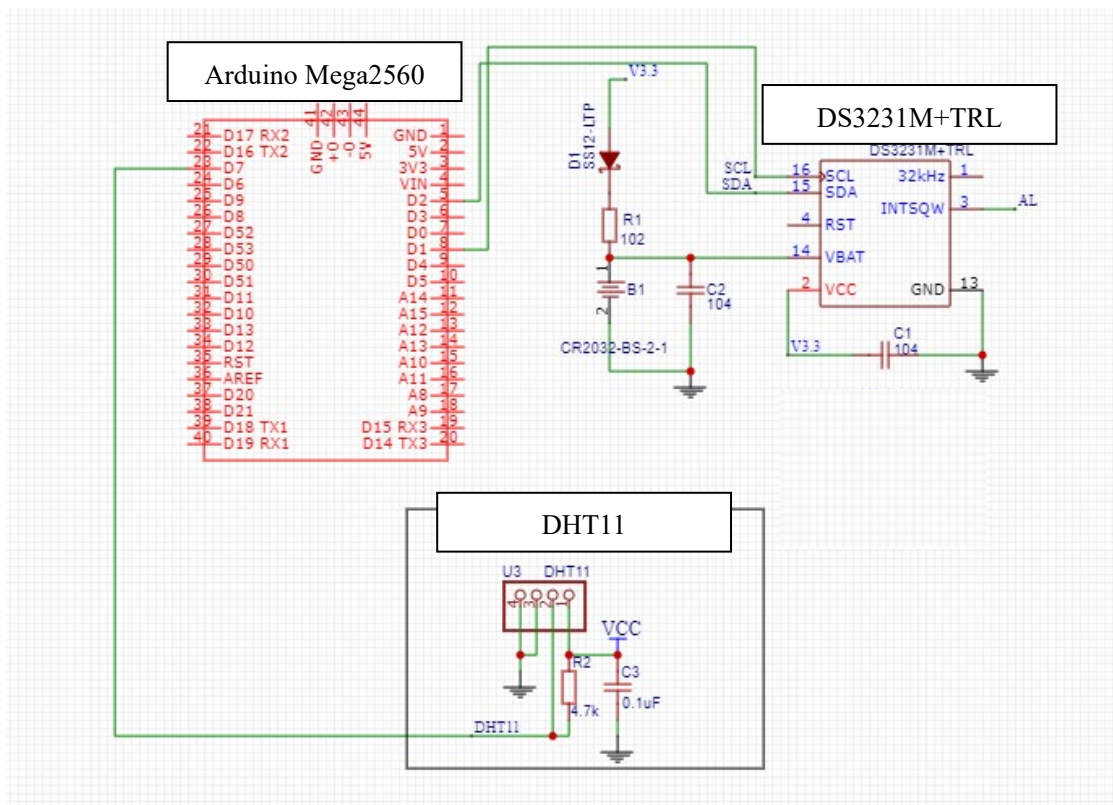


图 3-2 电子控制系统电路图

物联网模块基于 Arduino NANO 与 esp8266-01s。Arduino NANO 与 esp8266 之间通过 UART 通信，波特率 115200。Arduino NANO 向 esp8266 发送相应的 AT 指令，进行其初始化与 WIFI 连接，从而实现物联网功能。此方案为物联网功能实现的经典方案。

3.1.4 高级传感器系统

高级传感器系统分为非特定性人声识别与人体感应两个部分，两部分间无直接通信，且两部分分别与电子控制系统单向通信。

非特定性人声识别采用 REC-V2 模块，最高能够识别 1200 句不同的指令。模块具有完善的音频流控制系统，可以自动选择从外界环境或者 AUX 输入中采样，进而能够实现手机通话的远程控制。模块与电子控制系统之间通过 UART 通信，波特率 9600，单向传输识别结果。

人体感应由树莓派搭载 Ubuntu 系统，运行 OpenCV 处理摄像头采集的图像，每 10 帧执行一次如图 3-3 所示的数字图像处理，在保证识别效果的同时照顾了树莓派的较低运算能力。而后通过波特率 115200 的 UART 与底盘系统进行单向通信，完成人脸对中的操作，提高物联网管家的交互水平。数字图像处理的关键代码如图 3-4 所示。

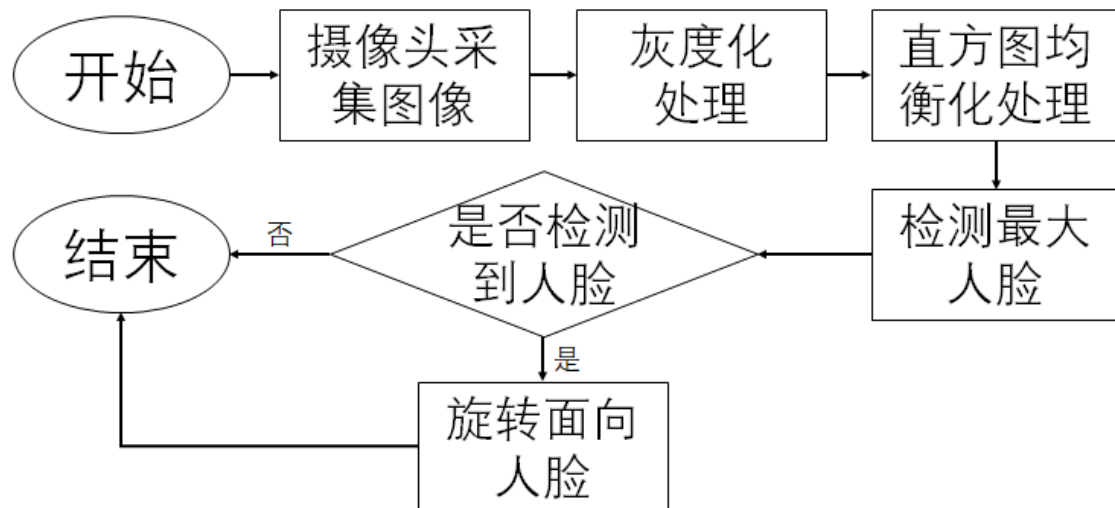


图 3-3 数字图像处理程序框图

//控制旋转

```

namespace _gpio {
    void init(void);
    inline void turnLeft(void);
    inline void turnRight(void);
    inline void stop(void);
    inline void find(void);
    inline void miss(void);
}
  
```

//人脸检测器

CascadeClassifier facedetect;

```

int main(void) {
    while (1) {
        camera >> camerafra; ++fcnt;
        //每10帧进行一次人脸检测
        if (fcnt == 20) {
            fcnt = 0;
            graylized(camerafra, grayedimg);           //灰度化处理
            equalizeHist(grayedimg, equalizedimg);     //直方图均衡化处理
            //人脸检测
            facedetect.detectMultiScale(equalizedimg, faces, searscalefac, minneigh, flag, minsize);
            Rect maxface; int s = 0;
            FindtheMaxFace(s, maxface);                //寻找最大人脸
            if (s > 0) {                                //找到最大人脸
                if (maxface.x <= 40) _gpio::turnLeft();
                else if (maxface.x >= 410) _gpio::turnRight();
                else _gpio::stop();
                _gpio::find();
            }
        }
    }
}
  
```

```

    }
    else {_gpio::stop();_gpio::miss();} //未找到最大人脸
  }
}
}

```

图 3-4 数字图像处理关键代码

3.2 系统机械结构设计

以上各元件的长宽高如表 3-1 所示。

表 3-1 各元件的长宽高数据

元件名称	长/mm	宽/mm	高/mm
L293D 电机驱动板和 Arduino UNO 合并元件	25	50	20
直流电机	35	20	20
充电宝	119	35	10
Arduino Mega	100	25	20
图像识别模块 (树莓派)	25	30	20
语音模块	30	30	20

综合考虑以上各系统的尺寸及联系，设计出了智能管家的机械结构。各元件的安装层数如爆炸图 3-5 所示，各元件的实际安装位置如图 3-6 所示。

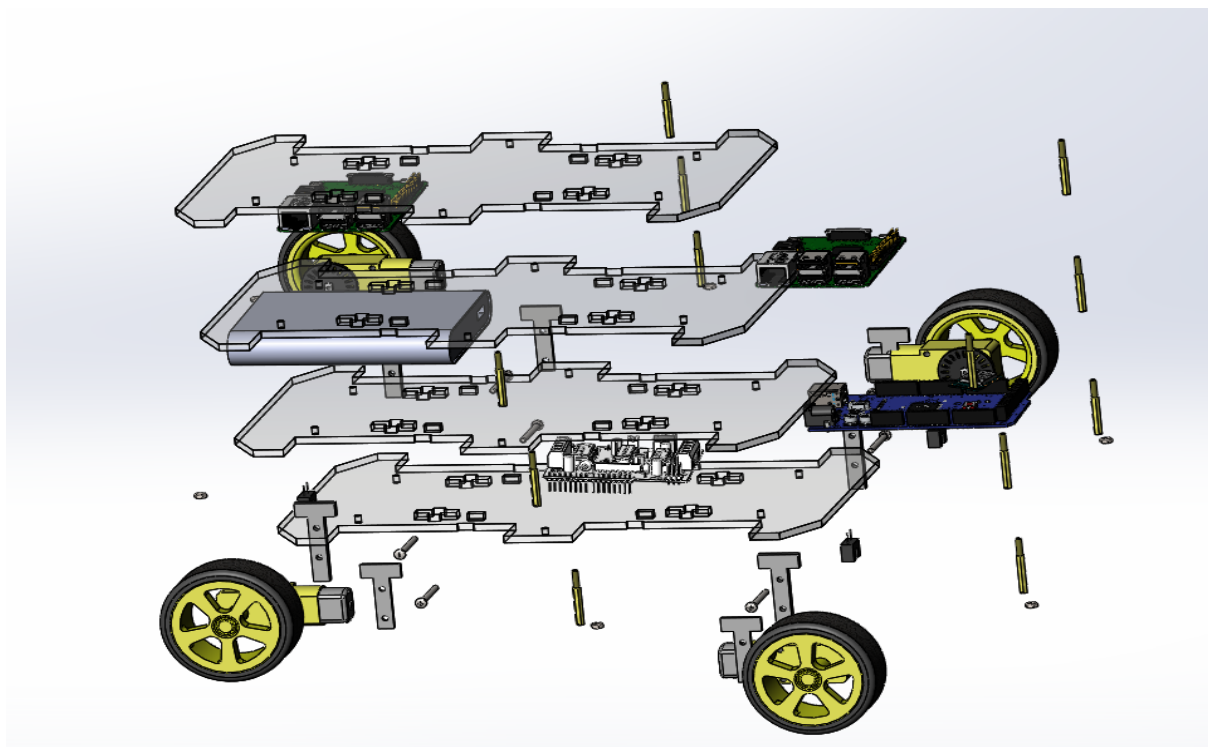


图 3-5 机械结构爆炸图

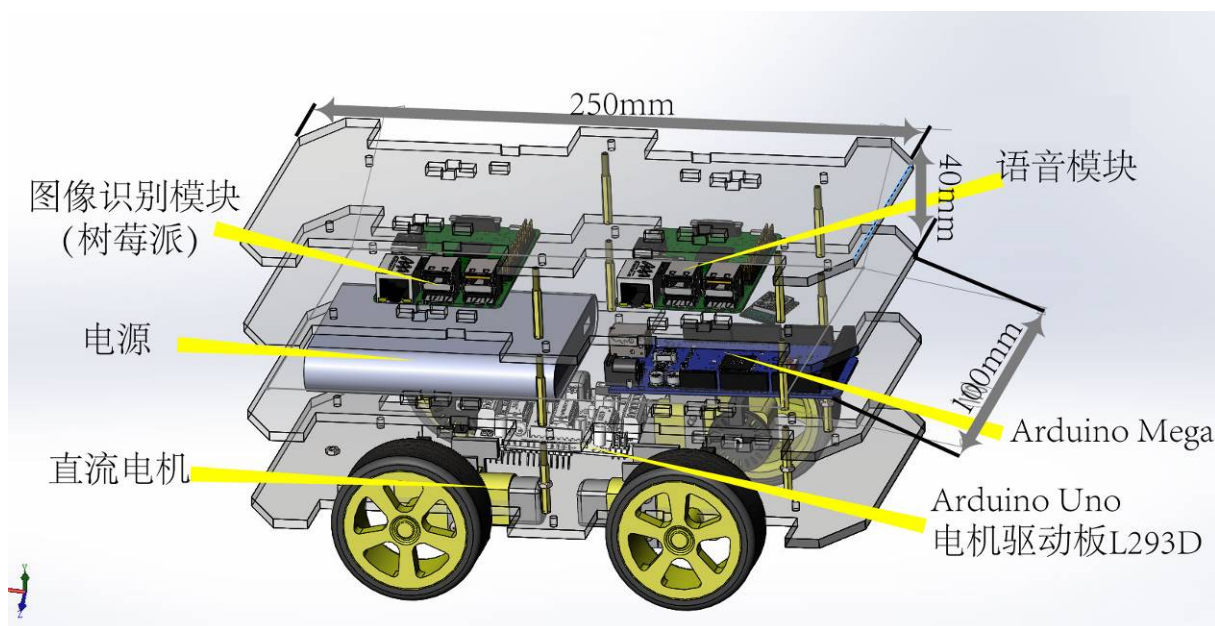


图 3-6 机械结构示意图

该系统以智能小车作为底盘，通过连接 4 块激光切割的亚克力板将小车层数增加至 3 层，从下至上依次为第一层、第二层和第三层。小车底盘长为 250mm，宽为 100mm，每层层高为 40mm。小车每一层的结构均为前后对称，即将所有元件固定在板的前后中心位置，其具体机械结构如下：

第一层：将 L293D 电机驱动板和 Arduino UNO 合并为一个元件并固定在第一层的正中心；在第一层的凹槽内分别安装麦克纳姆轮，并将直流电机固定在其内侧。L293D 电机驱动板和 Arduino UNO 合并元件和其两侧的直流电机距离均为 5mm。

第二层：将充电宝作为电源固定在第二层左侧，Arduino Mega 固定在第二层右侧，元件间距离 30mm 用于接线。充电宝和 Arduino Mega 与小车的左右两端均无间隔。

第三层：将图像识别模块（树莓派）固定在第三层距右侧 30mm 处，语音模块固定在第三层距左侧 55mm 处，且元件间距离 25mm 用于接线。

第四章 样机组装、测试与评估

4.1 系统组装

前文已经提到过，小车分三层，材质为亚克力板。为了提高组装的精确性，小组成员决定采用激光切割。小车实际车体如图 4-1 所示。

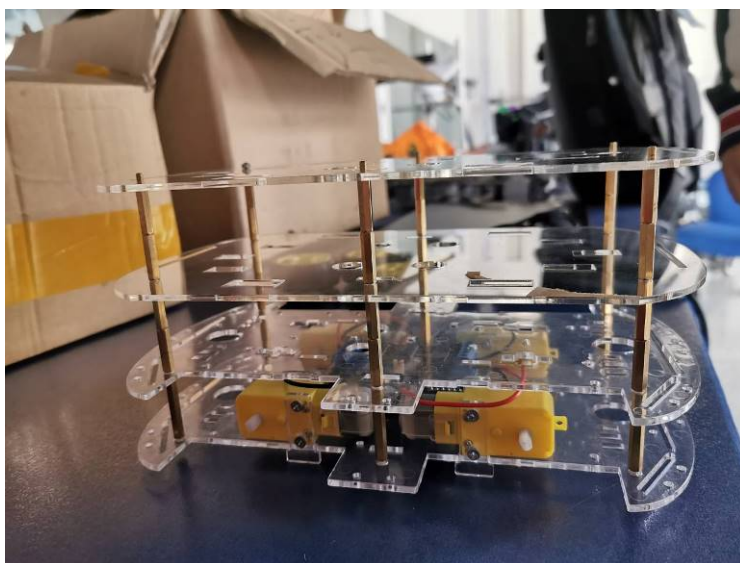


图 4-1 小车车体实物图

关于如何将多个电路板固定在小车上的问题，经过小组成员讨论之后，认为由于电路板自带螺丝孔，最好采用打孔并用螺丝固定的方式达成目的。经过对螺丝半径、长度的一番比较与小组成员的互相协作，成功将多个器件按照建模时的图形规划固定在了三层亚克力板上。最后制作出来的样机效果牢固整洁，后期对成品还准备进行进一步包装。

在打孔过程中，也遇到了打孔不齐的问题，经过对器件摆放方向的细微调整后，解决了这个问题。为了防止打孔过程中破坏亚克力板，小组成员找到了合适的受力面，选择了规格稍有不同的口径转头对不同的器件固定进行了打孔。

关于接线问题，经讨论决定，为了防止连线脱落，决定用热熔胶将其粘死，这是样机制作的最后步骤。

最终得到的样机如图 4-2 所示。

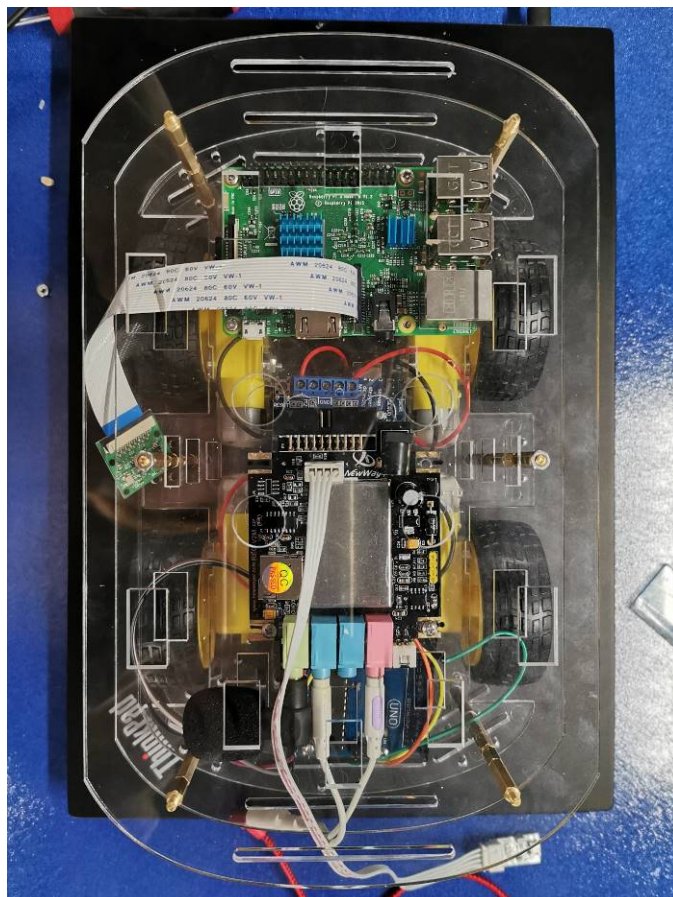


图 4-2 样机实物图

4.2 系统测试与评估

在组装好样机后，对样机进行了预定功能的测试。

在机械结构方面，该系统能够保证各部分元件的正常工作，且能够在桌面上进行平稳的移动。但由于该机械结构层数较多，重心较高，不适宜在较大的斜面上进行移动。

在电源方面，该样机在充电宝满电的情况下正常频率使用三天后仍能够供电，满足了预期续航三天以上的功能。

在各部分功能上，该样机能够正常显示时间温湿度，并能够进行动态表情的显示。该样机能够进行人脸的检测以及对应的旋转，也能够进行特定语音的互动。但由于系统储存空间有限，表情的数量有限，可以识别的语音数量有限，同时由于技术的限制，语音识别仍存在一定的误差。

第五章 项目总结

5.1 总结

总体看来，项目完成了预期的大部分功能，已经初步具备了演示能力，并具有一定的成本控制能力。然而截止至结项时仍存在部分问题。

5.1.1 表情交互单一

由于自身存储的数据较少，物联网管家的表情较为单一，且交互模式较为固定，难以给人一种亲切的感觉。

5.1.2 机械结构草率

样机整体是在市面上的智能小车的底盘的基础上采用激光切割亚克力增加层数，整体装配较为混乱，重心偏高，坚固程度差，体积过大，外形不具备感染力。

5.2 展望

5.2.1 削减成本

分布式结构导致样机整体的造价偏高，并伴随着大量的算力的浪费。下一代产品应当对整体的控制逻辑进行大幅度的修改，以便实现样机的商品化。

5.2.2 机电一体式设计

为了避免半数空间用于走线的尴尬局面，大型项目在设计时，通常会在装配图中考虑到实际的连线，甚至能够软件模拟电气性能。在实际操作中，缺少提前设计且复杂的 UART 通信线路造成了很多麻烦。因此，在下一代的制作中，我们希望能够在电路网络的设计中，尽可能多地使用数字化手段辅助设计，以得到最好的效果。

参考文献

- [1] Wei-Yu Chen, Frank Wu, and Chung-Chiang Hu. "Application of OpenCV in Asus Tinker Board for face recognition", Proc. SPIE 10443, Second International Workshop on Pattern Recognition, 104430J (19 June 2017).
- [2] 廖彬全, 罗佩, 马远佳. 基于智能语音交互系统的翻译机器人[J]. 信息与电脑(理论版), 2019, 31(17): 110-112.
- [3] 张水利, 吴瑞智, 李欢敏, 屈俊青. 一种具有语音功能的智能家用唤醒系统设计[J]. 微型电脑应用, 2018, 34(10): 1-3+8.
- [4] 酆涛. 基于人工智能的图像识别技术的研究[J]. 通讯世界, 2019, 26(08): 69-70.
- [5] 梁路宏, 艾海舟. 基于人脸检测的人脸跟踪算法[J]. 计算机工程与应用, 2001(17): 42-45.

谢辞

这个项目是我长期规划的一个项目，感谢上海交通大学给了我这个平台，让我结识了一些志同道合的小伙伴，并把这个项目，这个物联网管家付诸实践。同时交大为我们解决了后顾之忧，为我们提供了充足的经费，在这里再次感谢机械与动力工程学院。

感谢焦素娟老师课程前期的讲解与全程的细致的指导，感谢满孝臣助教长期的工作问题的解答。

感谢全班同学在长期的项目设计中的宝贵意见，这些意见为我们的设计节约了大量时间。

并对全部在项目制作过程中直接或间接提供帮助的老师同学表示衷心的感谢。