**韩 山 师 范 学 院**

学 生 毕 业 论 文

（2018届）

**题目（中文）基于树莓派智能镜子的设计与研究**

**（英文）Design and research of Smart Mirror Based on raspberry school**

**学院**： **物理与电子工程学院**

**专业： 电子信息工程 班级： 20181273班**

**姓名： 李铭皓 学号: 2018127345**

**指导教师： 傅胤荣 (副教授)**

目录

摘要

ABSTRACT

第一章:绪论

1.1 课题的研究目的和意义

1.2课题的研究现状

1.3课题的研究内容

1.4本章小结

第二章:树莓派概述

2.1 树莓派的定义

2.2树莓派使用的语言

2.3树莓派的发展方向

2.4树莓派的应用

2.5树莓派4b简介

2.6本章小结

第三章:硬件设计

3.1树莓派智能镜子基本结构

3.2树莓派智能镜子的系统工作原理

3.3本章小结

第四章:软件设计

4.1开发环境准备

4.1.1树莓派系统镜像的下载和烧写

# 4.1.2 OPENCV安装

# 4.1.3 Linux常用的语句

# 4.1.4网页爬虫API的申请

4.2树莓派智能镜子需要安装的库文件及安装方法

4.3程序设计

4.3.1启动摄像头并调用

4.3.2爬取网上信息

4.3.3设计画板

4.4本章小结

第五章:总结

参考文献

致谢

摘要

伴着科技的发展和社会的进步，智能家居在生活中所占领的重要性越来越大，而作为生活的必需品—镜子，也开始了走向智能化。智能镜子在2012年移动设备世博会中出现在人们的视野中，但由于智能镜子的技术不够成熟，以及所运用的器件比较昂贵，因此成本居高不下，所以导致了用户市场利用率低，为了改变这一现状，增加智能镜子的功能，对此做了研究与设计，首先运用了树莓派为核心，使用python语言进行编写程序，再通过python爬虫技术，和摄像头的调用，在显示屏上显示出来，从而实现日期、天气和穿衣指数的显示，并且能推送合适的衣物和购买途径，完成一个树莓派智能镜子。

关键词：智能镜子，智能，树莓派，python

ABSTRACT

With the development of science and technology and the progress of society, the importance of smart homes in life has become more and more important, and as a necessities of life-mirrors, they have also begun to become intelligent. Smart mirrors appeared in people’s field of vision at the 2012 World Expo for Mobile Devices. However, due to the immature technology of smart mirrors and the relatively expensive devices used, the cost remains high, resulting in low user market utilization. In order to change In this situation, the function of smart mirror has been added. Research and design have been done. First, Raspberry Pi is used as the core, and the program is written in python language, and then it is displayed on the display screen through the python crawler technology and the call of the camera. , So as to realize the display of the date, weather and clothing index, and can push the appropriate clothing and the way to buy, complete a Raspberry Pi smart mirror.

Keywords: smart mirror, smart, raspberry pie, python.

第一章：绪论

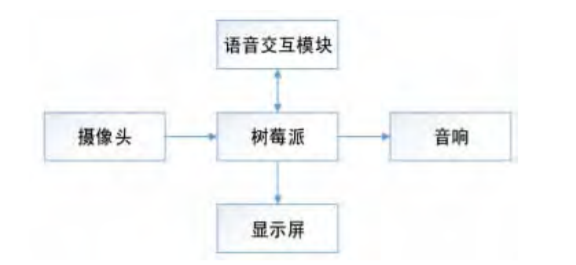
智能镜子是一个可以让你在照镜子的同时，知道天气，新闻信息，便签通知等信息的一个智能家居产品。在2012年一次移动设备世博会中被推出,之前的智能镜子是用安卓系统来运作的,主要用的部件有安卓平板电脑。智能镜子其实是一面集显示屏,天气监控,外接设备和信息交互等功能合为一体的镜子。被开启时，它就能变成一个智能的互助中心，可以变成用户在生活或者工作上的助手,家用智能镜子的显示器通常可以显示日常的天气、热点新闻，路况信息，日程安排，穿衣指数，健康管理等。智能镜子由于用到许多技术和较好的显示屏,所以价格一直很高,而用树莓派的话就只需要用python语言来编程,外接一个显示屏,这样的话就大大降低了成本和产品生产的难度,而且可以通过树莓派的wi-fi实时更新信息。

1.1本课题的研究目的和意义

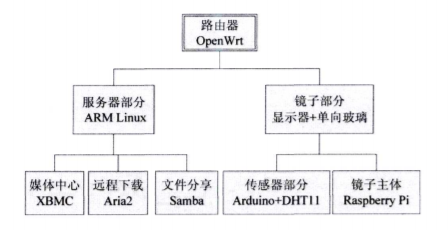
根据部分信息统计出来,女性和男性每日照镜子的次数约为23和16次, 可以说镜子是每家每户的必需品,但是单一的镜子已经不能满足现代人快节奏的要求,有时候用户会在照镜子的时候想知道一些额外的信息,比如今天的天气,新闻,当日的行程之类的,这时候就要整理妆容之后,然后再去查找想知道的信息,为了解放双手,节省更加多的时间,因此就有了智能镜子的诞生,所以智能镜子在智能家居这个领域里发展潜力还是很大的。

1.2课题的研究现状

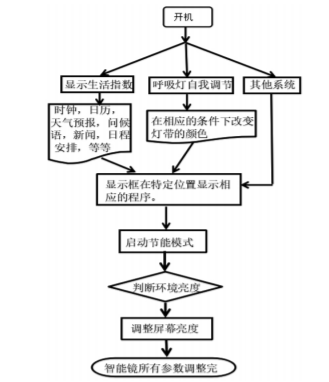
由张志浩、武 壮、赵海川、杨文凯编写的《内蒙古科技与经济》中基于树莓派的智能镜子研究与设计一章中,里面也有写到树莓派做智能镜子，他运用的是树莓派做核心，运用摄像头做镜子，然后通过api进行爬虫获取天气信息，他还有一个语音交互模块，可以进行对魔镜发出指令，再通过它的语音播报数据，智能镜子的主要功能都实现了，优点很明显：比正常的智能镜子加了语音交互模块和音响，从而实现与智能镜子的互动，当然，缺点也很明显：它只有天气数据，没有其他的数据显示，商业方面用途太少了。



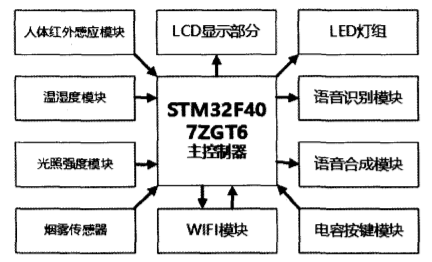
由亢文韬、洪智慧、沈鑫、李迎春编写《现代计算机(专业版)》这本书中的其中一章--基于ARM芯片的智能镜子的研究与实现，他这个智能镜子的底层操作系统是基于Linux和Windows的，它能实现天气预报、温湿度的检测、室内电器的遥控开关、社会媒体信息通知、储存服务器、媒体中心等功能，主要用到ARM主板、arduino控制版、显示屏、传感器元件和单向透视玻璃。路由器上运行着开源路由器固件OpenWrt，服务器操作系统需要ARM的Linux环境，使用Samba进行文件共享，显示界面运行媒体中心系统XBMC，Linux系统运行Apache、PHP和MySQL。他这个智能镜子优点好多，功能也十分强大，但显示那部分他使用的是显示屏和单向玻璃结合形成，这会使智能镜子显得十分笨重，同时他使用的语言没有python方便，这是它的缺点。



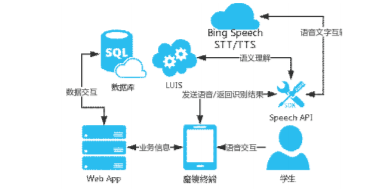
由赵凌峰、梁毓豪编写的《电子世界》中的：基于树莓派的智能镜子也说到智能镜子，他的设计主要用到显示屏、原子镜、树莓派、摄像头、麦克风模块。他是用原子镜和显示屏的搭配来充当镜子的效果，摄像头负责热红外人体感应，通过麦克风模块进行交互模式，外加了一个led灯作为呼吸灯，led由光照传感器和控制器模块进行控制。智能镜子应该有的功能都已经实现了，使用的也是python来编写，但镜子效果那是由原子镜和显示屏搭配而产生的，这会导致智能镜子有一种笨重感。



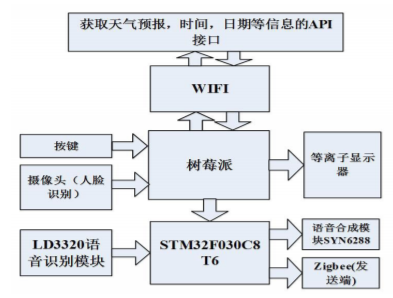
由肖培、胡天立、曹源、祝蓓、李璇编写的《科技创新与应用》中的基于STM32F407ZGT6的智能镜子设计，他这个设计的系统主控制器为STM32F407ZGT6，运用人体红外感应判断，然后控制显示器开关，利用温湿度模块获取温湿度，通过光照强度传感器控制led亮度，用烟雾传感器实现报警功能，还运用到ASR M08--A语音识别模块和语音播报 的语 音识别模块和语音合成模块和Wi-Fi模块进行信息交换。他们这个设计功能很多，缺点也很明显，他实现镜子的反射功能没用到摄像头也没用到单向原子镜，导致反射得不是很好，而且他运用stm32来制作，需要外加的模块太多了，编写起来没有python容易、简便。



由周志寰、蒋舒林、陶文龙编写的《科技创新导报》中一章基于Azure云计算平台的智能镜子的研究与实现也说到了智能镜子，他们是运用微软Azure认知服务设计的，里面还搭载了android8.0系统的ARM设备。在网络信息方面主要使用了Azure的AzreWeb App服务，Azure MySql服务和微软的BingSpeechAPI认知服务等微软云平台提供的服务。通过Android8.0系统的ARM设备进行语音交互，运用麦克风阵列来采集语音的信息，提供给应用程序进行处理，完成使用者指令的采集和对环境的感知。缺点是：需要使用的和调用的服务太多，编写的语音不单一，既要使用Android又要使用微软Azure，没有python简单、方便。



由单柯、郭楚滨、刘占宇、陆煜程、潘双龙编写的《科技传播》这本期刊中，有一章说到基于物联网的智能镜子，他的设计以树莓派为主控制，STM32为辅助控制，树莓派连接摄像头进行人脸识别，通过WiFi模块获取天气预报、时间、日期等信息，然后把这些信息放进去等离子显示器（单向透视镜加显示屏）上，还用到STM32来控制语音识别模块，从而实现交互功能。缺点：使用了单向透视镜，会使智能镜子太笨重，里面还用到stm32，这提高了成本，还提高了编写难度，建议可以使用树莓派进行语音识别和图像显示。



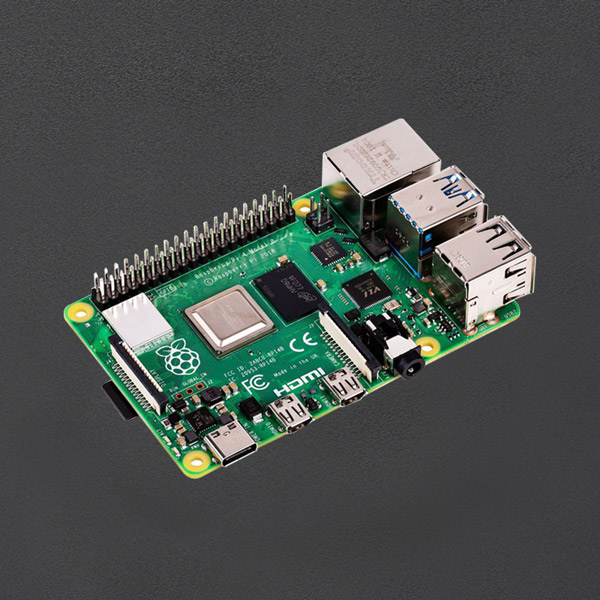
1.3课题的研究内容

现在的智能镜子要求,本次设计将以微型计算机为核心,和opencv打开摄像头，用户的基本信息,time库来调用出现在的时间,thread多线程库让它多个程序可以同时运行,由于里面需要的信息可能需要调用到网上的资料,因此我们还需要json库进行解析网页,和requests库请求网,还需要。首先设计出一个窗口,再通过python爬虫技术爬取网上的信息放进界面里面去,最后把界面优化就形成一个智能镜子。

第二章 树莓派概述

树莓派是由Eben·Upton/埃本·阿普顿作为项目带头人在英国的“Raspberry Pi 慈善基金会”开发。2012年3月，由埃本·阿普顿发售，外形微小，但包含了电脑的所有的基础功能，因此诞生了树莓派电脑板。

2.1 树莓派的定义



Raspberry Pi（也被称为"树莓派"）,里面装载了一个微电脑处理器,一个网口插座用来连接网线以及USB转接口,可用于连接鼠标、键盘，U盘传输数据；还有视频接口和音频接口，高清多媒体接口，电源接口，运用SD快闪存储器来存储数据，它可支持的操作系统有：基于ARM的Linux操作系统、树莓派媒体中心、RISC OS Pi、Raspberry Pi的复古游戏、FreeBSD、Chromium OS、Windows 10 IoT核心版、Android和Android Things、AROS：AmigaOS Remake、IchigoJam BASIC RPi。

Raspbian

Raspbian操作系统是官方推荐的操作系统，它是Debian系统在ARM上的编译版，再加上专门针对树莓派深度而定制的硬件驱动与软件程序。

Ubuntu Mate

Ubuntu Mate操作系统是Linux发行版本，Mate是Gnome 2桌面环境的衍生，曾作为Ubuntu的缺省桌面。Ubuntu Mate合适刚入手树莓派的同学使用，可支持的东西十分多，而且界面是这么多个系统里面最好看的，但是缺点就是没有官方的系统--Raspbian系统的CPU优化做得好。

Snappy Ubuntu Core

Snappy Ubuntu Core 操作系统是Ubuntu的其中一个版本，它是一个面对智能设备的平台，可运行依赖于云端里面或者存储在本地文件库里的相同软件，Snappy Ubuntu Core 系统最大的优点就是可以不用定期更新。

OpenELEC

OpenELEC操作系统是一个用户体验好、且运行快的一个XBMC媒体中心。里面提供家庭影院PC端所需要的各种软件，也可以使用在树莓派身上。

Pidora

Pidora操作系统虽然不是Fedora的官方版本，但Fedora在官网里推荐树莓派使用。Pidora系统的桌面环境是XFCE。

Windows 10 IoT

Windows 10 IoT系统它支持树莓派使用，Windows 10 IoT系统是可以给开发人员研究的。

RISC OS

RISC OS操作系统是一个由桌面和命令指示界面组成的系统。

Arch Linux ARM

Arch Linux操作系统是在ARM的架构上进行移植的，十分简洁，不自带软件，只有一个命令指示界面，对初学者不是很友好。

Kali Linux

Kali Linux操作系统是基于Debian系统而创作出来的Linux版，主要用于进行渗透测试和数字的取证，Kali Linux操作系统里面包含了许多进行透测的试软件，使用者可以通过liveCD、liveUSB和硬盘来运行Kali Linux操作系统。

2.2树莓派使用的语言

树莓派系统是基于Linux的，支持可编写语言有：python、java、c等，而树莓派的主要编写语言是python语言。Python语言不仅是一种功能强大并而完善的通用型语言，它同时也是一种解释型、弱类型、面向对象的脚本语言。相比于其他编程语言,Python编写的代码非常地简单，非常容易上手。比如有时候我们要实现某个功能，如果使用Java语言来编写的话，可能需要100行代码，但使用Python语言来编写的话，可能只需要20行代码就完成了，这是Python语言其中一个非常吸引人的特点。

Python语言的两大特色是可扩展性和清晰的语法：

1、Python语言的可扩展性非常大，主要在它的模块中显示出来，Python语言相对于其它的语言，它的脚本语言更加强大和丰富的类库（这些类库被人们称为“batteries included”- -内置电池），这些类库在许多的应用场景中都覆盖了，比如：数据库访问、网络编程、文本操作、GUI、文件I/O等。

2、Python语言编写的语法十分清晰，它的编写格式非常的严格。例如：当你使用到了if语句，那么你编写下一行的时候必须向右进行缩进，否则就会报错，不能编译。

此外，Python语言的讨论社区也非常发达，可能在一些很少见到的应用场景中，Python语言通常都有它相应对的开源模块用来提供解决的方案。

Python语言作为一种解释型的编程语言，它本生就配备了跨平台的能力，只要该平台有相对应的Python语言解释器，那么Python在该平台上运作是没有问题的。

Python语言的优点有：简单、开源、免费、容易维护、容易解读、容易上手、可使用的途径广泛、运行速度快、有较高可移植性、具有许多的库、编写代码规格严谨。

2.2树莓派的发展方向

树莓派发展趋势，体积小方便携带，高性能CPU处理，低价格化，电源端电流变小。

体积小方便携带：体积更加小，这样才能让人们随时随地携带，更好地使用和开发。

高性能CPU处理：只有高性能的CPU才能更加好地实现算法运算，使图像处理数据不会出现延迟性。

低价格化：只有价格降低，那么设计成本才能降低，才能更好地流入市场，用户的接受程度才会提高。

电源端电流变小：现在Raspberry Pi配置的是5V/2.5A电源，而生活中大多数都是5V/2A的配置，所以会导致高功耗，容易发热。

2.3树莓派的应用

树莓派可以实现的功能有:远程控制和自动备份、家庭防盗器或者监控器、DIY 街机、智能小车、机器人、人脸识别、无线路由器、家庭媒体中心等等，主要分为这几个：

1、在计算机网络和通信领域。

2、在智能家居领域。

3、在监视防盗领域。

4、机器视觉领域。

5、人工智能领域

2.4树莓派4b简介

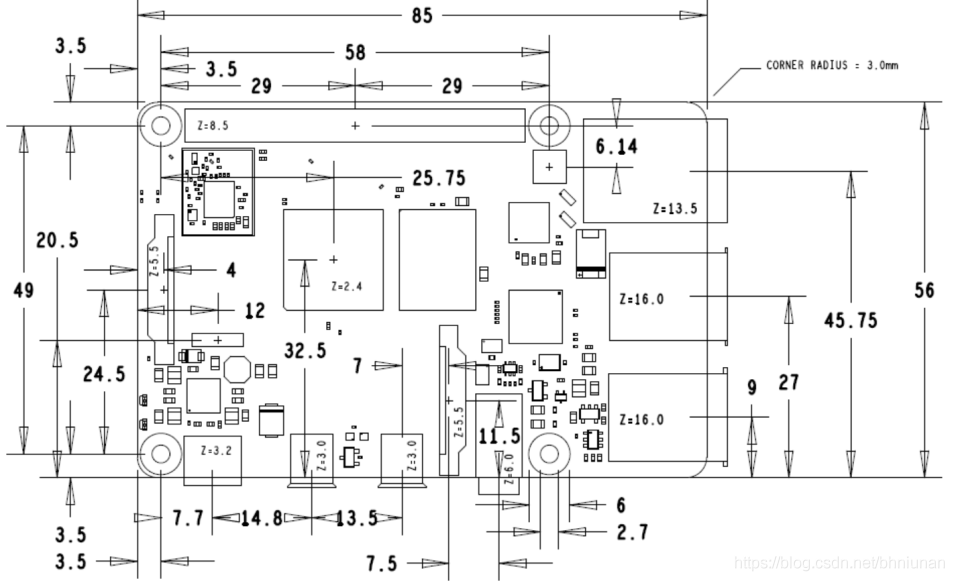
树莓派4b里面搭载了64位1.5GHz的四核处理器，使用的是VideoCore VI GPU，支持H.265、H.264、OpenGL ES3.0graphics，1G/2G/4G LPDDR4 内存，我使用的是4 G内存的，有PCI-E通道，支持蓝牙5.0 BLE，四个USB接口，可支持4K分辨率，一个USB Type-C接口。

树莓派4 B是使用型号为BCM2711的CPU，相对于之前树莓派3 B +的博通处理器，更加新的版本，树莓派4B处理器里面包含了更好的性能优化和更强的散热器。而且允许更加多的时钟频率，能更加准确地监察和控制芯片温度，使芯片不会温度过高。

Raspberry Pi 4B （树莓派4B）有更加多的带宽，这使树莓派4B能够提供千兆网速，而之前的树莓派型号最高只能达到330兆网速。

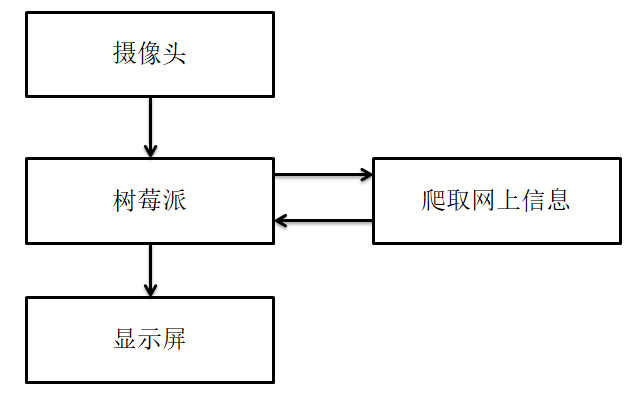
Raspberry Pi 4B （树莓派4B）的核心时钟是500 MHz的VideoCore VI,比以前 400 MHz的Broadcom VideoCore IV更加好。

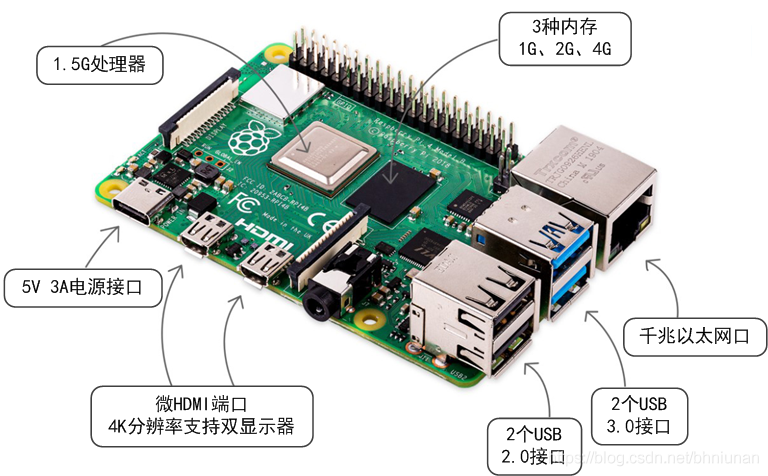




第三章:硬件设计

3.1树莓派智能镜子基本结构





智能镜子最主要使用到的是摄像头、Raspberry Pi（树莓派）、显示屏、SD快闪卡，和python语言爬虫技术，其中Raspberry Pi（树莓派）是本设计的核心部件，充当着“管家”的角色，操作系统是Linux，通过5V 3A的电源接口供电，显示屏是通过Raspberry Pi（树莓派）的USB端口进行通电，微HDMI端口进行图像传输，本设计的工作过程为：首先要把SD卡放进去microSD接口，通过命令指令窗口使能摄像头并且打开摄像头，并且通过程序把摄像头的显示通过自己想要的规格显示出来，然后把数据传输到显示屏，连接显示屏的是通过树莓派的USB3.0接口和HDMI端口进行连接，连接摄像头则是通过树莓派的CSI（摄像头）和DSI（显示屏）排线接口（在HDMI端口旁边的排线接口）进行连接。

3.2树莓派智能镜子的系统工作原理

基于现在的智能镜子要求,本次设计将以微型计算机为核心,tkinter窗口库作为界面显示用户的基本信息,time库来调用出现在的时间,thread多线程库让它多个程序可以同时运行,由于里面需要的信息可能需要调用到网上的资料,因此我们还需要json库进行解析网页,和requests库请求网页，还需要re库进行正则运算提取所需要的信息。首先设计出一个窗口,再通过python爬虫技术爬取网上的信息（天气、穿衣指数、服饰推荐），把这些信息全部放进画板界面里面去,通过程序打开摄像头，调整摄像头显示界面的大小、位置和分辨率，最后把画板界面进行优化一下，把信息进行合理的布局就形成一个智能镜子。

3.3本章小结

第四章:软件设计

4.1.1开发环境准备

树莓派没有配置FLASH，因为树莓派是通过SD卡启动，因此我们需要下载镜像文件，而且将镜像文件烧写到SD卡上，然后启动烧写的系统就可以用了。(这个镜像文件里面有：bootloader、kernel、文件系统)。

这个镜像文件里面有树莓派官方的raspbian系统和已经配置和修改好基础环境的raspbian系统。

因为树莓官方系统的镜像文件原始设定是没有开启远程服务的，这样会导致没有显示屏的使用者在使用树莓派的时候增加难度，为了针对这个问题，我安装的系统镜像文件是已经开启了远程服务的，当烧录好这款镜像系统文件后就可实现远程登录树莓派的系统了。

以下是获取已经开启了远程服务的镜像系统网站网址：https://www.yahboom.com/study/raspberry4B

提取的密码是：nzv8

将里面的文件下载到本地磁盘后进行解压，解压后你会得到一个img 文件。

这里的注意点是：

1、要使用专门的工具用来格式化内存卡，我使用的是SDFormatter。

2、安装一个windows系统下烧写镜像文件的工具，我使用的是：Win32DiskImager。

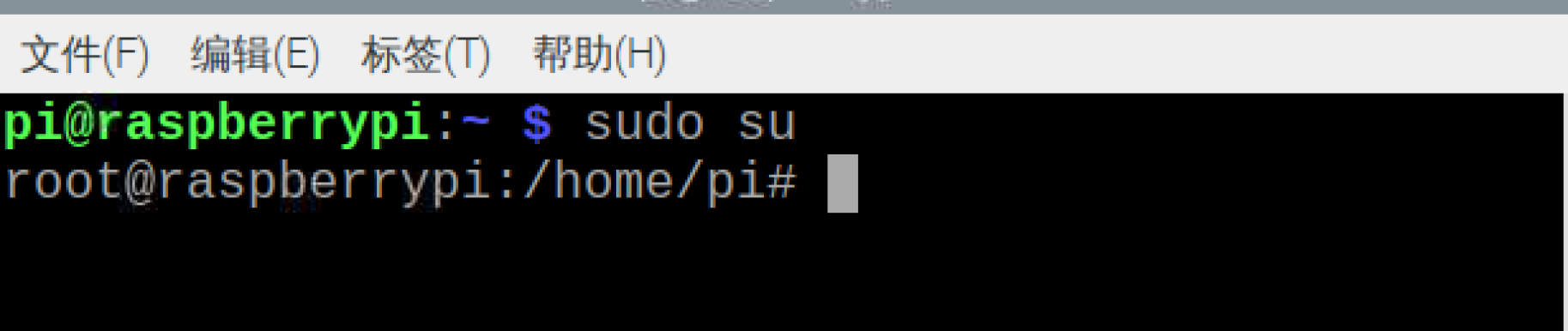
烧写镜像文件步骤：

打开Win32DiskImager，在Win32DiskImager软件中选择解压出来的镜像文件，在“Device”下选择SD的盘库，点击“Write”，然后就会自动开始烧写系统，安装的时间由你SD卡不同而不一样，如果你SD卡的运行速度快的话，时间会变短，当烧写完成后，会弹出烧写完成的对话框，就说明烧写已经完成了（如果烧写不成功，把后台的防火墙关闭，和一些控制防火墙的软件都关闭），我使用的是4G的SD卡，所以烧写完毕后，SD卡会显示里面的内存只有74MB，不用担心，这个是正常的现象， Windows系统是看不见linux的磁盘分区。

# 4.1.2 OPENCV安装

首先帮树莓派连上网络，打开命令指示窗口输入：

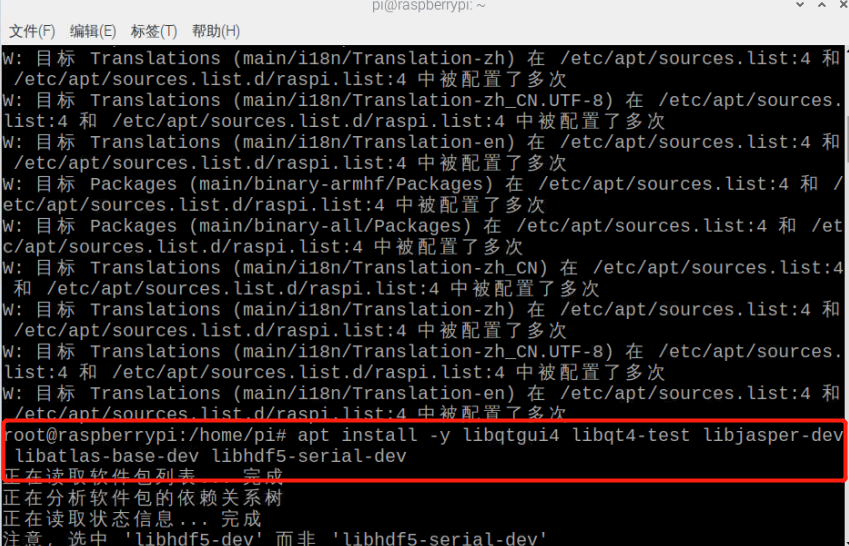
sudo su（切换成管理者模式）



再输入apt update更新系统

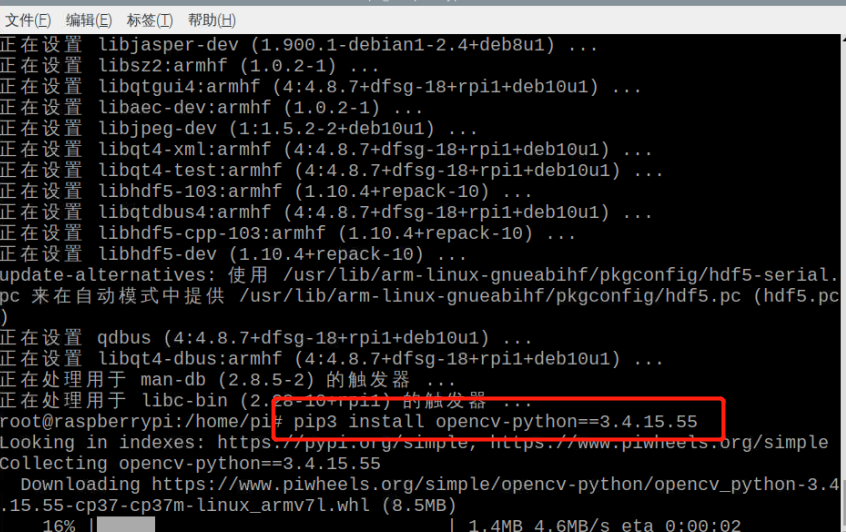


输入apt install -y libqtgui4 libqt4-test libjasper-dev libatlas-base-dev libhdf5-serial-dev安装依赖包

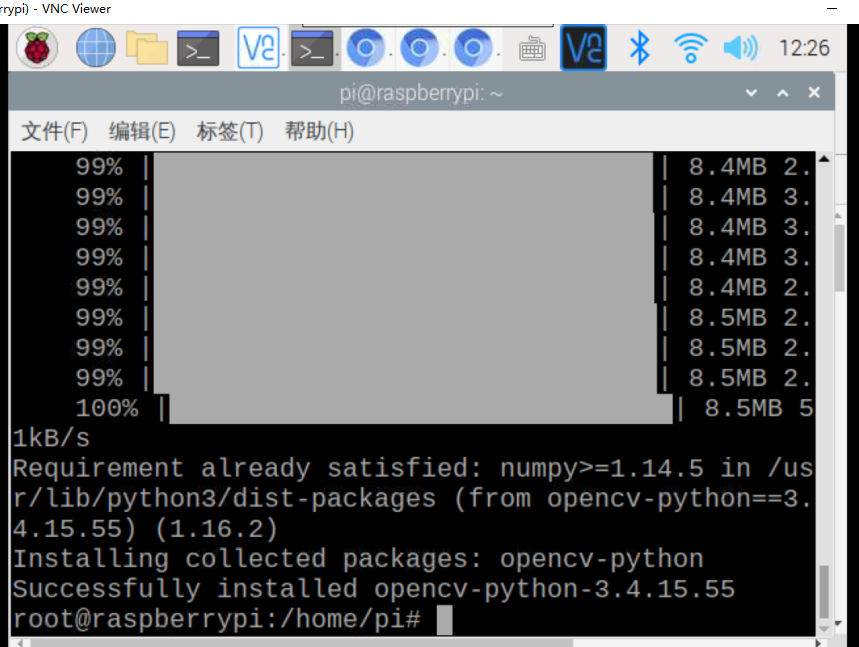


都安装好后就代表你已经完成了安装OPENCV的前期准备工作了。

接下来输入pip3 install opencv-python==3.4.15.55

****

完成之后是这样的：



接下来我们还需要安装OPENCV的函数和常用的模块，方便以后使用OPENCV，输入这句语句：

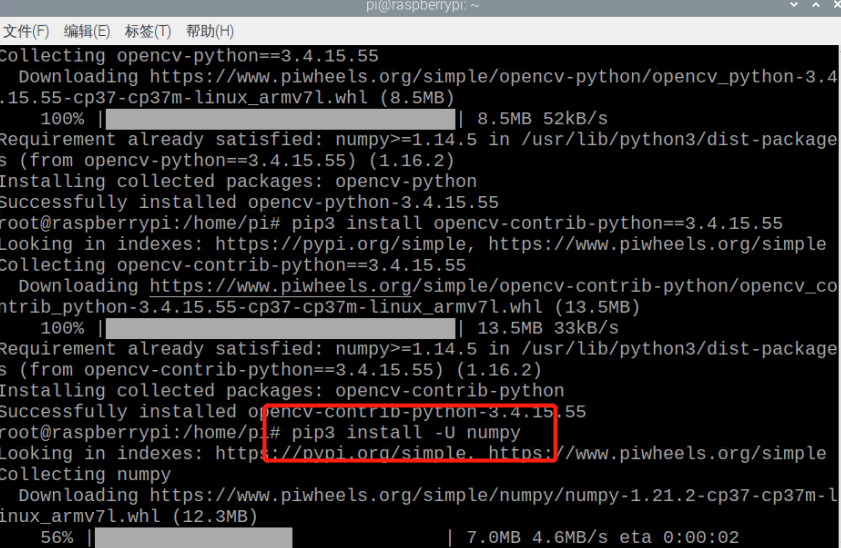
pip3 install opencv-contrib-python==3.4.15.55

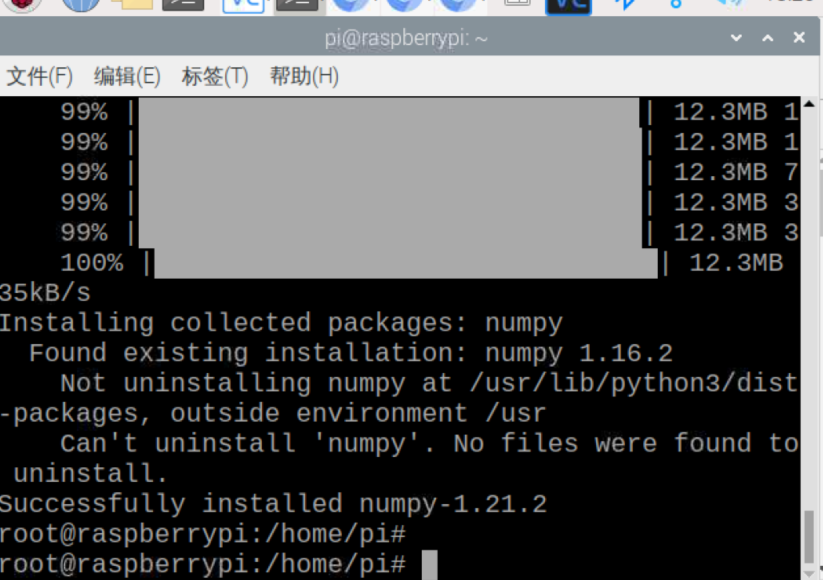


安装好后，由于我们使用的是python语言，所以还需要安装一个python包，我们在命令指示窗口输入：

pip3 install -U numpy

NumPy(Numerical Python)这个库是Python语言的其中一个扩展程序的库，NumPy库能够支持大量的矩阵和维度数组运算，同时它还能为数组的运算提供许多的数学函数库。





当显示上面这张图的界面的时候，证明你已经安装完成了。

# 4.1.3 使用树莓派时常用的Linux命令

# 我们常用Linux的命令有：

cat /proc/version 查看操作系统版本

cat /proc/cpuinfo 查看主板版本

df -h 查看SD 存储卡剩余空间

Ifconfig 查看ip 地址

tar –zcvf filename.tar.gz dirname 压缩

ar –zxvf filename.tar.gz 解压

linux 系统常用 apt(Advanced Package Tool)高级软件工具来安装软件

sudo apt-get install xxx 安装软件。

sudo apt-get update 更新软件列表。

sudo apt-get upgrade 更新已安装软件。

sudo apt-get remove xxx 删除软件

# 4.1.4网页爬虫API的申请

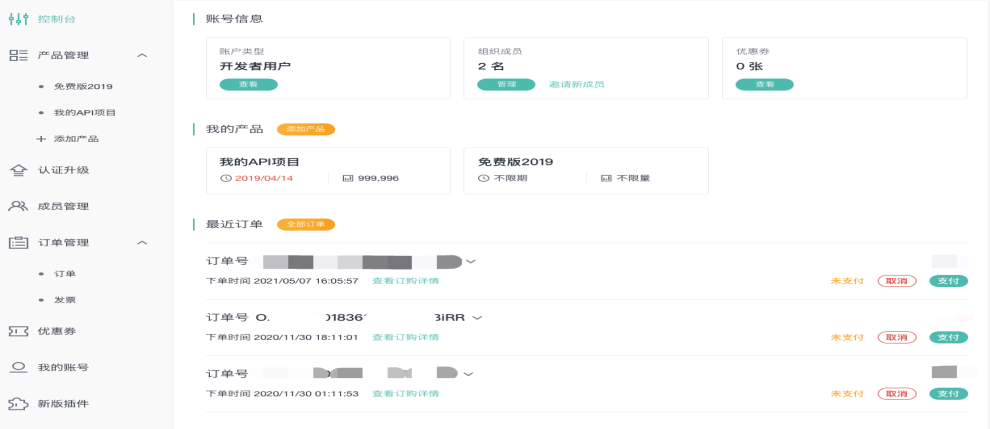
这里我使用的是心知天气的信息，心知天气是一家由中国气象局官方授权的商业气象服务公司,它基于气象数值的预报和人工智能技术,可以为我们提供高精度的气象数据、天气监控机器人、气象数据可视化产品,以及能源、电力、保险、农业、交通、互联网、物联网...

心知天气与中国气象局气象数据中心战略合作，而且接入了美国的NOAA和NASA以及欧洲的ECMWF、英国的MetOffice等全球多家数据源，再通过AI人工智能技术产生更加精准的天气数据。

使用心知天气我们首先要登录心知天气的官网注册一个账号，这样才能使用 API 接口。



注册成功后,就可以在控制台中看到您的账户类型,产品等信息,下图就是控制台：



注册成功选用免费版就好了，然后进入控制台：



进入控制台后，你可以点击你想要获取的信息，你就能获取到相对应的产品的公钥和私钥（即API），这里我选择的是天气。



不同数据的API接口请求文档在对应的数据页面中



通过网址填入你的API和你想要查找的地点就会给予你相应的信息，我获取的是潮州的天气：

var = {"results": [{"location": {"id": "WS4YTUBPDZY9", "name": "潮州", "country": "CN", "path": "潮州,潮州,广东,中国",  
 "timezone": "Asia/Shanghai", "timezone\_offset": "+08:00"},  
 "now": {"text": "多云", "code": "4", "temperature": "31"},  
 "last\_update": "2021-10-16T15:53:23+08:00"}]}

它给予你的是一个字典，你想要的信息全部都在这个字典里面，接下来就需要你通过字典的调用就好了。

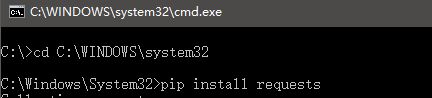
4.2树莓派智能镜子需要使用的库文件及安装方法

制作树莓派智能镜子需要用到许多的库文件，有些文件是python自带的，有一些文件则需要自己额外安装

需要安装requests库

Requests库是python实现的简单易用的HTTP库，使用起来比urllib简洁很多由于Requests库是第三方库，因此在使用前我们需要cmd（命令指示窗口）进行安装，我们在命令指示窗口输入：

pip install requests



安装完成后，当你编写代码的时候import requests一下，显示正常的话，则说明可以开始使用了。

Requests库是用于Python编写的，它是基于urllib，而且采用Apache2 Licensed开源协议的HTTP库；

相比urllib库，Requests库更加方便，它可以节省我们许多的工作和精力，Requests库它完全满足HTTP的测试需求，Requests库经常也被用来进行接口测试。（用了requests库之后，你基本都不太愿意使用urllib库了）用一句话来说，requests库就是用python来实现的目前为止最简单、方便、易用的HTTP库，所以通常都建议你想运用爬虫技术的话，那就推荐使用requests库。

使用到json库

同样我们需要cmd（命令指示窗口）进行安装，我们在命令窗口输入：

pip install json

**json库是用来处理json格式数据的Python标准库，json库里面主要包括两种函数，分别是：解析函数与操作函数**。

（1）解析函数：主要完成解析键值对的内容。

（2）操作函数：主要用于外部json格式数据和在程序内部数据各种类型之间进行转换的功能。

（3）json格式包含了对象与数组，分别用方括号[]与大括号{}进行表示，分别对应键值对的对等关系与组合关系，在使用json库时，我们需要留意json格式中 “对象”与“数组”的概念和Python语言中的“字典”与“列表”两者之间的区别和联系。

（4）一般来说，json格式中的内容将会被json库进行解析成为字典，json格式中的数组也将会被解析成为列表。

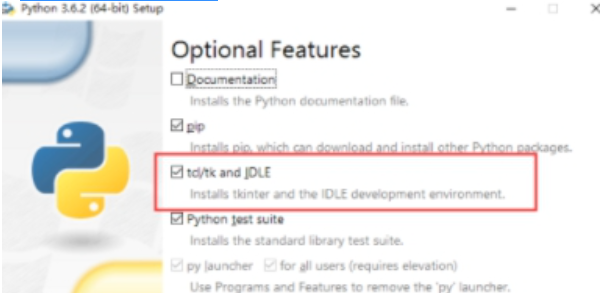
使用到time库

time库是Python语言中用于处理时间的标准基础库，它有着提供和获取系统的时间并且格式化输出的功能，它还提供系统级别的精确计时功能，主要用于程序的性能分析。time模块主要提供各种用于操作时间的函数，time库通常有两种方式来表示时间的: 第一种是用时间戳的方式(比如相对于2021.11.23 01:01:01以秒来计算的偏移量),这时候时间戳具有惟一性。第二种用数组的形式来表示，即(struct\_time),一共有九个不同的元素，它们分别表示为，同一个时间戳的struct\_time数组会因为时区不同而改变。

安装Tkinter库

Tkinter库是python中本身就自带的库，通常我们是不需要去安装的，可以直接就拿来使用，如果有些同学不清楚自己是否已经安装成功Tkinter库，那么就需要学习怎么去查看Tkinter库自己安装是否成功了，下面是给大家详细整理了查看是否安装好Tkinter库的步骤：

首先需要打开命令指示窗口，在里面打开python的交互界面，输入import tkinter，看是否有报错，没有报错就说明你已经安装好了，如果有报错的话，那就重新安装python就好了，在安装的时候勾选tk选项就可以了。



Tkinter库是Python的标准的GUI库。Python 使用 Tkinter库可以很快地创建GUI应用程序。由于Tkinter库是内置到python里面的安装包,所以只需要安装好Python之后就能 import Tkinter 库、而且 IDLE 也是用 Tkinter 编写而成、对于简单的图形界面 Tkinter 还是能应付自如。

使用threading库

threading库是用于线性程的，线程是由操作系统可以直接支持的执行单元，所以，高级语言常常都需要内置多线程进行支持，当然,Python也不例外，并且，Python的线程是真正的Posix Thread，而不是模拟出来的线程。

Python的标准库提供了两个模块：一个是\_thread模块和另一个是threading模块，\_thread是比较低级的模块，而threading是高级的模块，它对\_thread模块进行了封装。在绝大多数的情况下，我们只需要使用threading这个高级的模块就可以了。

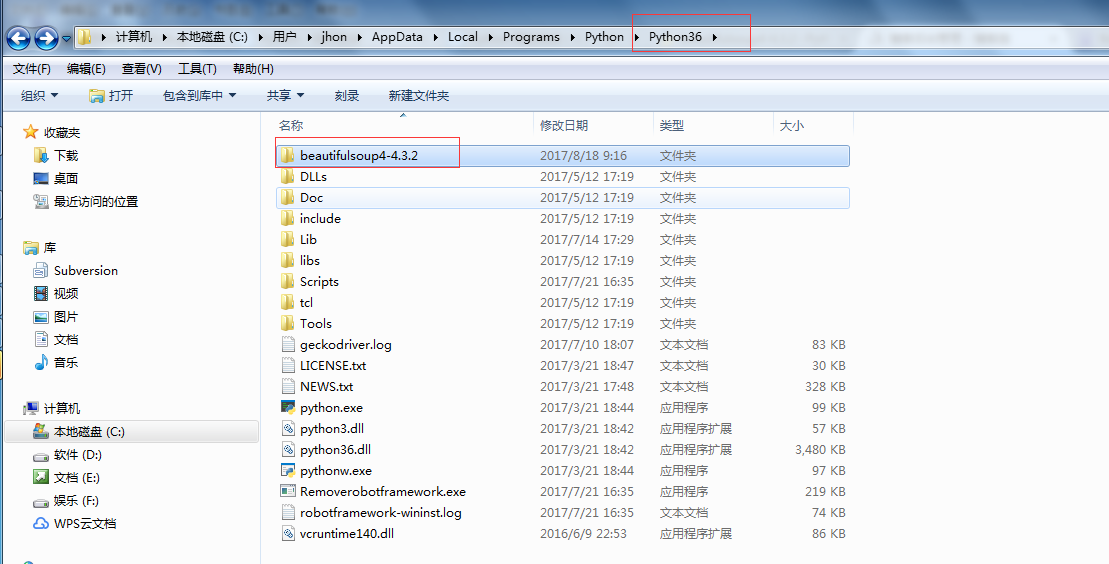
安装Beautiful Soup库

Beautiful Soup库是用python语言写的一个关于HTML/XML的解析器，Beautiful Soup库可以处理那些不规范的标记并且生成parse tree (剖析树)。它能为我们提供既常用又简单的navigating(导航)、修改剖析树以及search（搜索）的操作。它可以为我们大大节省编程的时间。Beautiful Soup库可以通过解析文档来获取网页文档里面所需要的数据，它能使用户很方便地从XHTML或HTML的文档中提取出数据，作为python的一个辅助工作，也是爬虫实践中的好帮手。

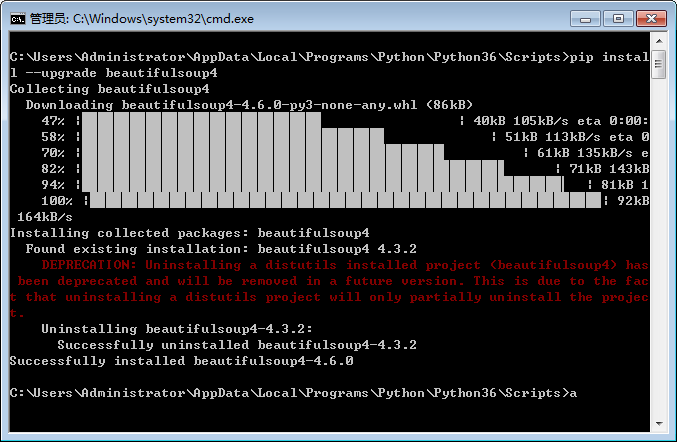
安装Beautiful Soup库的方法有两种：一种是使用pip命令进行安装，一种是直接下载压缩包安装。

第一种使用pip的命令安装步骤：打开命令指示窗口，输入：pip install beautifulsoup4就可以开始安装了。如果你找错了PIP的安装路径或者不知道PIP的安装路径，窗口弹出了：‘unknow or unsurrpotted command 'install'’这样的语句，显示错误，你可以输入命令：‘where pip’，它就会弹出你PIP的具体目录。

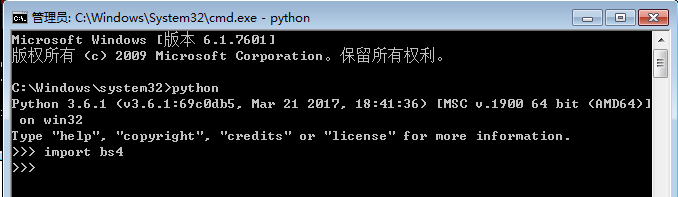
第二下载压缩包安装步骤：首先在浏览器中输入官方网址：https://pypi.python.org/pypi/beautifulsoup4/4.3.2进行下载压缩包，然后把压缩包解压，再将解压的文件夹放在python的文件夹下，如图：



然后进入beautifulsoup4-4.3.2这个文件夹下，打开命令指示窗口输入：setup.py build，再输入下面这条命令：setup.py install，最后验证一下beautifulsoup4-4.3.2是否已经安装成功，输入：pythonimport bs4，如果没有显示报错就是说成功安装，如果报出了以下的错误提示:'cannot import name 'HTMLParseError'', 其实报错的原因是安装的beautifulsoup版本太低，需要进行升级，因此我们只需要升级下beautifulsoup4-4.3.2就行了，我们可以使用PIP命令来升级，步骤：打开Python下PIP的安装路径，打开命令指示窗口，输入命令：pip install --upgrade beautifulsoup4，如图所示：



升级成功后，关闭命令指示窗口，再次进行验证，输入python import bs4，然后如图所示：



那说明你已经安装成功了。

使用re模块

Re模块是python中的正则表达式，正则表达式是用于处理字符串最常用的方法，我们在编码中随处可见正则表达式的身影。正则表达式通常都是大同小异，python语言中的正则表达式与其他的语言相比，有少许不一样：1、在替换字符串的时候，替换的字符串可以是一个函数。2、split函数可以用来指定分割的次数。3、前项界定的表达式必须定长。

通过使用正则表达式，它的功能：

可以测试字符串内的模式。——比如：可以用于测试输入的字符串，查看输入的字符串内是否有出现电话号码模式或者信用卡号码模式，这个模式也被称为数据验证。

替换文本。——可以用正则表达式来识别出文档中选定的文本，然后把该文本完全删除或者使用其他文本来替换它。

基于模式匹配从字符串中提取子字符串。——可以用于查找文档内或输入特定区域内的文本。

有两种工作模式，分别是：非贪婪模式和贪婪模式

正则表达式经常用于查找匹配的字符串。Python语言里面的数量词默认是贪婪模式，它总是会尽可能多的寻找相匹配的字符；非贪婪模式与贪婪模式是相反的，它总是会尽可能少的寻找相匹配的字符。

在python语言默认是使用贪婪模式，如果要使用非贪婪模式的话，只需要在量词的后面加上一个问号“?”就可以切换成非贪婪模式。

例子abbb：使用贪婪模式，我想要找“ab”的字符串，如果里面已经找到了含有“ab”字符串时，那它就表示为已经表达式匹配成功，但因为使用的是贪婪模式，因此还会继续往后匹配，一直匹配到最后一个是“b”，直到后面没有可以进行匹配的字符串的时候，匹配才会结束。然后返回匹配结果“abbb”。所以，可以将贪婪模式理解为：在整个表达式匹配成功的前提下，尽可能多的匹配。

非贪婪模式就是将我们例子中的正则表达式“ab+”改成“ab+?”，如果找到相互匹配的“ab”时，一旦已经成功匹配，就立刻结束匹配，不会继续往后面进行尝试匹配，然后返回成功匹配的字符串“ab”。

因此，非贪婪模式可以理解为：当整个表达式成功匹配的前提下，就会结束匹配，尽可能少的匹配。

正则表达式的对象有：

re.RegexObject

re.compile()：用于返回RegexObject的对象。

re.MatchObject

group()：用于返回被re匹配到的字符串。

start()：用于返回开始匹配的位置

end()：用于返回结束匹配的位置

span()：用于返回一个元组包含匹配(开始与结束)的位置

使用**urllib**库

**urllib库是python里面的标准库，简单来说就是安装python的时候已经自带的库，想要使用的时候只需要import调用一下就好。**

Urllib库是python语言中的一个用于操作URL、功能强大、并且在运用爬虫的时候经常需要用到的库。在Python2里面，被分为Urllib库与Urllib2库，Python3之后全部都被合并到Urllib库中，所以使用方法也有一些不同。

**urllib库里面包含了4个模块：**

**request模块：是最基础的HTTP请求模块，用于发起请求信号，就像访问网页，需要在浏览器上输入网址一样。**

**Error模块：是异常处理模块，当请求出现了错误，就需要使用这个模块来抓住异常，这样才能保证程序不会因为出现异常的时候而停止运作。**

**Parse模块：是一个工具模块，它提供了很多URL处理的方法，比如：URL的拆分、合并等等。**

**Robotparser模块：是用来识别目标网站里面的robot.txt文件（很少用到）。**

**可以使用request模块发送请求，方法主要有两个：urlopen() Request()**和**urlopen()。**