

基于百度 AI 与 STM32 的人脸、语音与物体识别系统研究

杨焕峥, 杨国华, 徐 玲

(无锡商业职业技术学院, 江苏无锡 214153)

【摘要】文章依托百度开放人工智能技术, 基于 STM32F407 芯片, 开发了包括人脸、语音与物体识别在内的本地人工智能识别系统。该系统通过 STM32F407 驱动 OV2640 摄像头拍摄人脸或物体照片, 调用百度 REST API 给予的标准 HTTP 端口, 按照 JSON 数据格式, 通过 88w8782 WiFi 模块将数据上传百度 AI 云平台进行识别, 接收文字识别结果, 并通过百度 AI 语音合成接口获得一段 MP3 进行语音播放, 实现人工智能识别技术的实际应用。

【关键词】百度人工智能; STM32F407; 语音识别; 物体识别

【doi:10.3969/j.issn.2095-7661.2018.04.009】

【中图分类号】TP391.4

【文献标识码】A

【文章编号】2095-7661(2018)04-0028-03

Research on face, speech and object recognition system based on Baidu AI and STM32

YANG Huan-zheng, YANG Guo-hua, XU Ling

(Wuxi Institute of Commerce, Wuxi, Jiangsu, China 214153)

Abstract: Based on the open AI technology of Baidu and STM32F407 chip, this paper develops a local AI recognition system including face, speech and object recognition. The system uses STM32F407 to drive OV2640 camera to take pictures of faces or objects, calls the standard HTTP port given by Baidu REST API, uploads data to Baidu AI cloud platform for recognition through 88w8782 WiFi module according to JSON data format, receives the result of character recognition, and obtains a piece of MP3 through Baidu AI voice synthesis interface to make voice play and realize the practical application of artificial intelligence recognition technology.

Keywords: Baidu artificial intelligence; STM32F407; speech recognition; object recognition

计算机对于人工智能识别一般要进行复杂的深度学习算法运算和神经网络模型的训练等过程, 而且准确度要高、耗时要短, 对计算机硬件配置要求较高, 甚至要由计算机集群来实现^[1]。由普通的卡片或个人计算机等构成的不依赖于网络的嵌入式人工智能识别系统在准确性和耗时上有欠缺, 而且训练模型需要大量信息数据, 个人或小型企业较难独立完成复杂的人工智能识别系统的搭建。人工智能识别系统一般包括联网的云端和本地两部分, 在允许依赖网络的情况下, 由云端数据中心的大规模的计算去完成人工智能识别算法, 本地实现具体的智能控制等应用。百度采用深度学习等算法, 以 API 形式开放人工智能技术^[2],

研制基于 STM32F407 的本地人工智能识别平台, 与百度 AI 云平台通过 WiFi 进行数据通信, 完成包括人脸、语音与物体识别在内的应用系统。

1 系统结构和功能

基于 STM32F407 的本地人工智能识别平台结构包括 STM32F407 MCU 控制器、OV2640 摄像头电路、88w8782 WiFi 电路、PCM1754DBQ 音频电路、W25Q-80VSSIG FLASH 电路、CP2102 UART 转 USB 电路、EZ1086CM-3.3 电源电路等组成, 连接 WiFi 路由器经过 Internet 网络和百度 AI 识别云平台进行 WiFi 通信, 如图 1 所示。

【收稿日期】 2018-08-16

【作者简介】 杨焕峥(1980-), 男, 江苏无锡人, 无锡商业职业技术学院讲师, 高级技师, 工程硕士, 研究方向: 嵌入式系统工程。

【基金项目】 2018 年江苏省大学生创新创业训练计划项目“百度 AI 语音识别在嵌入式系统中的应用”(项目编号: 201812702017X); 2016 年江苏省教育厅高校自然科学基金项目“工业过程时滞系统频域辨识方法研究”(项目编号: 16KJB120007)。

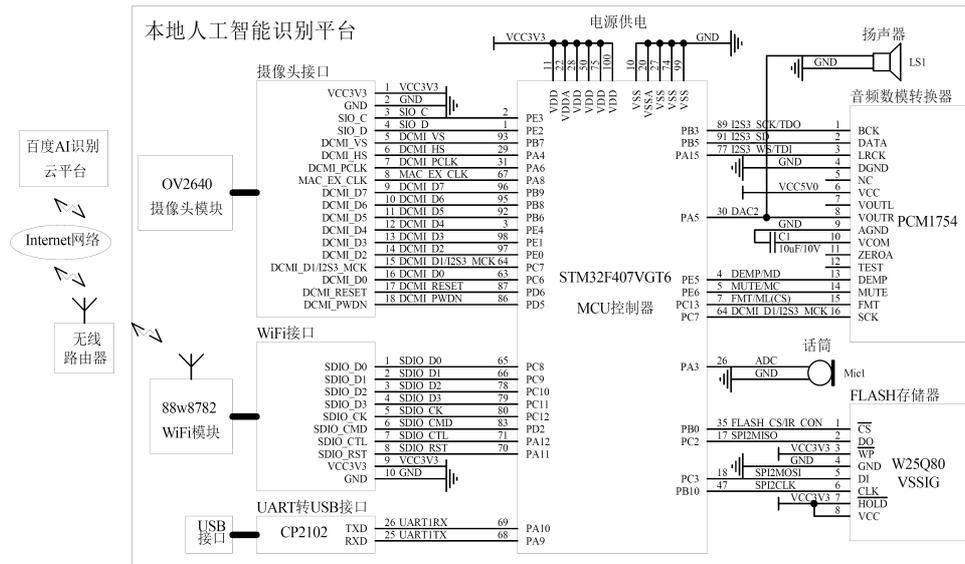


图 1 本地 AI 识别平台主要硬件电路图

OV2640 是 OmniVision 公司生产的 1/4 寸的 CMOS UXGA(1632*1232)图像传感器。STM32F407 芯片有一个数字摄像头接口 DCMI，使能系统时钟，配置相关引脚的复用功能，设置 DCMI 工作模式及 PCLK/HSYNC/VSYSNC 等参数，设置 DMA，启动 DCMI 捕获图像数据。

2 本地 AI 识别平台软件系统

2.1 软件系统结构

STM32F407 MCU 通过芯片驱动控制自带的功能部件，含 SDIO、UART、SPI 驱动等，通过外设驱动控制连接的板级外部硬件，含 88w8782 WiFi、pcm1754 驱动、Audio 等，移植 $\mu C/OS-II$ 内核，运行小型操作系统，实现多任务实时处理，通过 OS 相关 API，如任务、信号、事件、定时器等满足应用程序所需的功能，如语音对讲、MP3 播放等，跟 WiFi 模块之间通过 802.11 协议接口通信，通过 WiFi 相关 API 和移植的 LwIP 协议栈进行数据收发，通过 Socket 相关 API，如 connect、close、send、recv 等满足应用程序所需的功能^[3]，如图 2 所示。

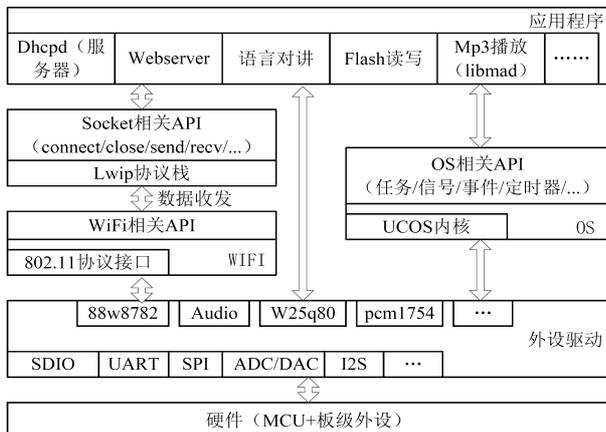


图 2 本地 AI 识别平台软件系统结构图

2.2 使用 $\mu C/OS-II$ 内核、LwIP 与 WiFi 驱动

系统需要具有外网访问能力，即设备需要连接到路由器 WiFi 热点^[4]，使用 $\mu C/OS-II$ 的工程。采用 C 语言编写 STM32F407 MCU 系统代码，通过 Keil for ARM 软件编译代码，将路由器 WiFi 热点的名字和密码写入编写的 default_web_cfg() 函数中，设备将会通过 DHCP 自动获得 IP 地址。特殊情况下如果想 WiFi 网卡或者以太网卡使用静态的 IP 地址，可以在代码里设置。设备支持基于百度人工智能平台的图像以及语音类 AI 应用，可在代码 app_cfg.h 中设置。其中，SUPPORT_AUDIO 表示使能百度 AI 智能语音平台应用，这个应用可以识别人的声音，也可以实现文字到语音转换。SUPPORT_CAM 表示使能百度 AI 人脸识别应用，可以识别人脸，并实现简单设备控制。使用这个应用前，需要先将人脸信息注册到百度 AI 平台。

2.3 编写 OV2640 摄像头驱动

摄像头驱动初始化过程。在文件 camera.c 中，编写 open_camera() 函数，初始化 DCMI，包括 IO 口和 DMA 的初始化，分配 DMA 缓冲区，缓冲区采用环形工作方式，所以必须及时读取图片，否则会被新图片覆盖。初始化 OV2640，通过 I2C 接口初始化 2640 寄存器，2640 的时钟是 STM32 提供的，创建 img_send_thread 线程用来发送图片到远程端。

图片采集和发送过程。在文件 camera.c 中，编写 DCMI_IRQHandler() 函数，判断 DCMI_IT_FRAME 标志，并从 DMA 缓冲区读取一个图片，封装成一个消息发给 img_send_thread() 函数。编写 img_send_thread() 函数，等待 img_queue 消息，接收到 img_queue 事件，取出数据，将图片发送到远程端，比如局域网或发送到 MQTT 服务器。

2.4 STM32F407 软件系统初始化

编写系统上电及所有模块的初始化过程程序。

在文件 `startup_stm32f4xx.s` 中, 加载主堆栈指针 `__initial_sp`, 供 `main` 和中断使用; 加载 `Reset_Handler`, 进入 `Reset_Handler` 复位使用。编写 `SystemInit()` 函数, 初始化 `pll` 与中断向量。

在文件 `main.c` 中, 编写 `main()` 函数, 初始化 OS 部件 `OSInit`; 初始化内存动态分配管理 `_mem_init`; 创建 `main_thread`; 开始线程调度 `OSStart`, 此时还没有 `sys_tick`, OS 会选择优先级最高的 `main_thread` 运行, 开始线程调度之后 PC 指针将不会回到 `main()` 函数。编写 `main_thread()` 函数, 初始化一些底层驱动的杂项, 包括 IO 口配置, `systick` 配置, 微秒级定时器的初始化等; 初始化按键和指示灯 IO 口; 初始化 OS 状态, 用于统计 CPU 使用率; 初始化 1、3 两个串口, 用于调试信息输出; 初始化片外 `spi flash`, `flash` 主要用于保存配置信息, 包括通过 Web 配置的 IP 地址路由器名称、密码等信息; 初始化音频驱动, 主要是 ADC 和 DAC; 初始化工作线程, 主要是定时器; 加载 Web 配置信息, 从 `spi flash` 读取; 创建 GUI 测试线程, 创建一些窗口; 初始化 LwIP 协议栈; 初始化 SDIO 和 WiFi 驱动; 初始化 `monitor` 功能, 创建 `monitor` 线程, 用于打印系统信息; 初始化 LwIP 底层与 WiFi 相关的部分; 初始化以太网口驱动, 创建一个线程用于处理以太网口中断, 以及在线、掉线检测; 初始化应用层的一些杂项, 创建一个定时器用于控制指示灯的闪烁, 添加一些应用初始化; 打开音频, 开启 ADC 和 DAC 转换; 初始化按键驱动, 创建一个定时器用于按键检测; 初始化实时时钟; 初始化一些传感器, 并创建一个定时器, 每 3 秒采集一次传感器数据; 初始化摄像头, 并创建一个线程用于图片发送; 初始化 USB 驱动, 并创建一个定时器用于 U 盘插拔检测; 进入主工作例程。

在文件 `work.c` 中, 编写 `work_process()` 函数, 初始化一些跟远程控制相关的部件, 比如 MQTT; 发送切换到 `station` 的事件给自己; 开始等待接收事件, 后续的操作都通过事件来驱动, 没有事件时就等待。程序之前发送的一个 `MODE_DETECT_CHANGE_STA` 事件, 这里会立即处理, 具体处理内容就是切换 WiFi 至 `STATION` 模式, 连接到路由器, 为后面的 MQTT 连接和远程控制做准备。编写 `init_iot_work()` 函数, 初始化 `iot_work` 结构变量; 将 WiFi 的 MAC 地址通过 MD5 算法生成一个 32 byte 的字符串, 这个字符串将用于 MQTT 订阅和发布主题; 初始化 MQTT, 创建一个线程用于 MQTT 连接和数据收发; 创建一个线程用于 TCP/IP 数据接收, 通过 `select` 机制接收 TCP 和 UDP 数据, 凡是通过 `add_select_array` 加入到 `select_fd` 数组的 `socket` 都会通过这个线程接收数据, MQTT 数据接

收则由 `mqtt lib` 来处理。

3 百度 AI 云平台

百度 AI 平台提供语音识别与合成、人脸识别、物体识别等技术服务, 只要以 HTTPS 的方式向平台请求服务, 就可以得到 AI 分析结果。百度 AI 平台调用开放 API 需要使用 OAuth2.0 授权, 需要在 URL 里包含 `AccesssToken` 参数, 可以向授权服务地址发送请求来获取。

3.1 百度 AI 平台接入——人脸识别

在使用人脸识别功能之前, 需要在百度平台注册, 创建应用, 并管理人脸库。注册并登陆百度云平台, 在控制台上选择人脸识别, 然后创建应用^[9]。创建的应用默认只能访问人脸识别相关的 API, 手动选择百度语音和图像识别相关的 API 接口, 后面的其它 AI 功能可以复用这个应用。创建好应用后, 百度 AI 平台会给应用分配好 API Key 和 Secret Key。接下来, 进行人脸库管理, 需要将用户的人脸信息录入到应用中, 作为后续人脸认证的参考。在应用的人脸库中新建组, 并将组名填写为 `test_group`。之后在这个组中创建用户, 并上传用户人脸照片, 这里的人脸照越清晰, 后续识别的准确度越高。创建用户名可填写为 `test_user`。把应用的 API Key 和 Secret Key 填写到 STM32F407 MCU 代码对应的宏中, 同时, 将人脸识别应用对应的宏打开, 编译并下载代码到设备。人脸识别应用初步实现为一个人脸识别控制继电器开关的功能, 测试时, 将设备摄像头对准人脸, 一般间隔 50cm 左右, 如果设备检测到百度人脸库里的信息和当前人脸相匹配, 则会控制继电器闭合, 否则使继电器断开。能够在继电器上接入电子锁来实现识别人脸然后控制电子锁开关功能。

3.2 百度 AI 平台接入——语音识别 / 语音合成

语音识别、语音合成与人脸识别使用同一个 AI 应用, 有相同的 API Key 和 Secret Key, 创建应用并将对应的 API Key 和 Secret Key 填写到 STM32F407 MCU 代码中。修改宏开关, 打开语音应用相关的宏, 重新编译并下载代码到设备。

语音识别 / 合成的功能如下: 先按下设备上的录音按键, 设备启动捕捉 PCM 音频数据, 时间长度为 3 秒, 然后传输给百度 AI 平台, 由平台算法将音频数据识别成文字, 并下传至设备, 此时可以检查识别的文字里有没有“打开”、“关闭”字符串, 如有则控制继电器执行对应动作。语音合成的功能为, 设备将语音识别的文字传输至百度 AI 语音合成单元, 得到一段 MP3 并下传至设备进行播放^[9], 也可以使用通信串口向设备发送数据, 发送数据的编码使用 UTF-8, 发送数据最前面可以加数字来选择想使用的人声模式, 同

(下转第 36 页)

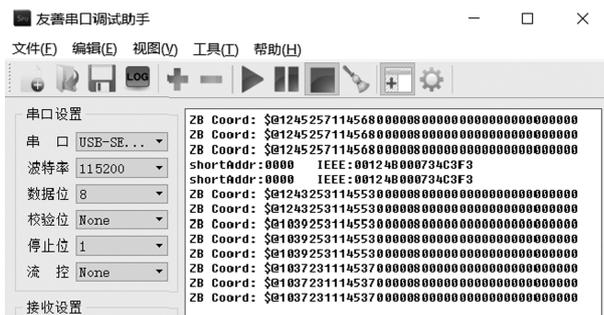


图 16 ZigBee 协调器串口发送数据图

5 总结

本文利用 CC2530 以及无线传感器网络技术设计了一种智能停车场的管理系统来解决停车难问题。在系统的实现过程中完成了 CC2530 核心板模块设计、协调器节点模块、终端节点模块、路由器节点模块软硬件设计,同时利用 ZigBee 无线网络技术初步建立停车管理系统,本系统的人机交互系统、远程客户预定系统还需进一步开发,以便尽早运用到商业停车场的智能停车场管理系统。

(上接第 30 页)

样传输至百度 AI 进行语音合成。

3.3 百度 AI 平台接入——物体识别

物体识别同时使用到了语音和图像功能,由于设备 STM32F407 MCU 资源的限制,需要在芯片内存空间使用上对应用代码做一些优化。

物体识别同前面的人脸识别使用同一个 AI 应用,有相同的 API Key 和 Secret Key。编译和下载代码到设备,可以对物体识别功能进行测试,设备使用摄像头周期性地捕捉图像并上传至百度 AI 平台进行物体识别,同时设备把接收到的百度 AI 文字识别结果再上传百度 AI 语音合成单元,合成为 MP3 接收并播放出来。测试时,设备播报“识别开始”后,设备会将正对准摄像头的待识别物品识别并播报出来。

4 结束语

研制基于 STM32F407 的本地人工智能平台,与百度 AI 云平台进行数据通信,完成包括人脸、语音与物体识别在内的应用系统。STM32F407 驱动 OV2640 摄像头拍摄人脸或物体照片,编写程序,调用百度 REST API 给予的标准 HTTP 端口,按照 JSON 数据格

【参考文献】

- [1]刘军.基于 ZigBee 的智能停车场管理系统的设计与实现[D].哈尔滨:哈尔滨工程大学,2012.
- [2]原景超,郭涛,王思宁.无线低功耗采煤机状态在线监测系统[J].煤炭工程,2018(8):18-20.
- [3]葛广英.ZigBee 原理、实践及综合应用[M].北京:清华大学出版社,2015.
- [4]罗庆生,韩宝玲.一种基于超声波与红外线探测技术的测距定位系统[J].计算机测量与控制,2005(4):304-306,334.
- [5]冯金龙,钱蕾,丁力.一种基于无线网络的 LED 智能照明系统的研究与设计[J].武汉职业技术学院学报,2018(3):93-97.
- [6]曾博,李英姿,冯家欢,张建华,刘宗歧.计及电动汽车无功支撑能力的分布式电源与智能停车场联合规划方法[J].电工技术学报,2017(23):185-197.
- [7]张衡.基于 J2EE 平台的智能停车系统设计与实现[D].大连:大连理工大学,2017.
- [8]戴明.一种智能无线路灯控制系统的设计[J].湖南邮电职业技术学院学报,2018(1):9-11,17.

式,通过 88w8782 WiFi 模块将数据上传百度 AI 云平台进行识别,接收文字识别结果,并通过百度 AI 语音合成接口获得一段 MP3 进行语音播放,实现人工智能识别技术的实际应用。

【参考文献】

- [1]Peng An, Danping Cao, Xiaoli Yang, et al.Research and application of reservoir classification method based on deep learning [C]//CPS/SEG Beijing 2018 International Geophysical Conference & Exposition Electronic papers,2018:1762-1765.
- [2]赵楠,缙珊珊.人工智能应用现状及关键技术研究[J].中国电子科学研究院学报,2017(6):590-592.
- [3]赵鹏.模块化嵌入式人工智能软件的开发[J].电子技术与软件工程,2018(4):49.
- [4]禹禄君.浅析移动互联网体系架构[J].湖南邮电职业技术学院学报,2014(3):59-61.
- [5]罗佳伟,孙雪峰,李琳.基于百度 AI 平台的 Web 人脸注册和登录系统的设计[J].中国新通信,2018(6):73.
- [6]郝雯超,冯月芹,李春光.基于嵌入式平台的实用语音识别研究[J].电子器件,2018(1):110-114.