

用虚谷号搭建文档高拍仪原型

罗亮 王盼盼 浙江省萧山中学

浙教版高中信息技术教材《数据与计算》，有一个答题卡填图识别算法解析示例，因为贴近生活，学生们都很感兴趣。这个案例可以进一步拓展，如通过文档高拍仪对答题卡数据进行采集，那就能实现一个真实的答题卡识别工具。当然，最好的拓展莫过于自己搭建一个功能更加强大的智能文档高拍仪原型，那样不仅可以识别答题卡，还可以识别文字、二维码之类。

● 文档高拍仪原型的功能分析

文档高拍仪的基本功能是图像采集，增强功能是光学字符识别、条码识别等。图像采集分辨率是决定高拍仪型号的重要依据，像素太高，对图片的处理速度和传输会提出新的要求，一般来说，500W像素比较合适。可见，文档高拍仪的原型在硬件方面要求很低，选择任何一款迷你电脑再结合摄像头即可。

在学校众多的开源硬件中，学生选择虚谷号作为核心硬件。虚谷号是一个面向人工智能教学和Python编程学习的开源硬件平台，板内集成高性能处理器和通用单

片机、兼容Arduino，可挂载摄像头等各种通用外设，可以外接键盘显示器成为独立主机，也可以自带电源独立运行Arduino、Python程序进行相应信息处理。想象一下，虚谷、充电宝、摄像头组合在一个外壳之中，形成手持设备，随走随拍，是多么方便的事情啊。

图像采集完成以后，可以尝试使用开源的光学字符识别软件进行字符识别或者使用百度大脑AI开放平台的图片文本识别功能。可供选择的基于Python的开源光学字符识别软件有darknet-ocr、

Pyocr、pytesseract、Tesseract OCR、CNOCR。这些开源软件各有擅长之处，CNOCR适合识别图片中的中文，但该开源软件需要在Windows平台下进行识别，所以笔者最终还是采用调用百度AI功能实现文档高拍仪的增强功能。

使用高拍仪肯定需要了解当前工作状态，虚谷号自带的LED灯则派上了用场。当然，也可以加一块小型的OLED显示屏。经过以上分析，可以得到文档高拍仪原型所需软硬件，如图1所示，对应处理流程如图2所示。

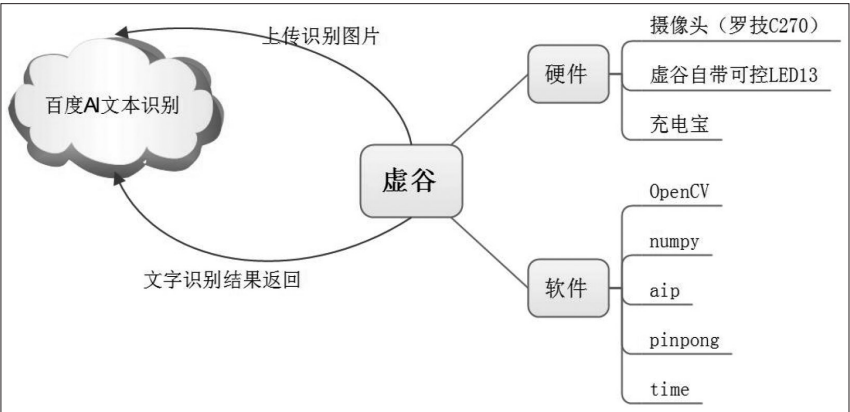


图1 文档高拍仪原型涉及软件、硬件模型

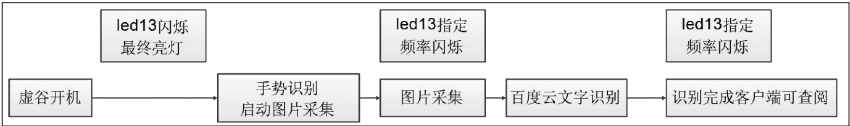


图2 文档高拍仪原型自动化工作流程

● 文档高拍仪原型功能实现

笔者采用模块化思想将高拍仪功能分解细化,逐一实现,然后合并功能,大致分成图片采集、启动处理、状态表示、OCR识别、结果呈现五大方面。

1. 图片采集

图片采集是视觉处理Opencv模块的一个基本功能。笔者将USB摄像头通过USB连接上虚谷号之后,就完成了硬件准备。frame代表摄像头采集的像素内容,后期需要对frame进行集中处理。在“测试场景”窗口为当前活动窗口的情况下,按'q'键退出图片采集,按's'键将采集的图像以文件形式保存。在后期功能合成时,如cv.imshow、cv.destroyAllWindows这些与目标功能达成不一致且耗费资源的代码可以略过不予使用。相关关键性程序代码如图3所示。

2. 启动图片采集识别

一般情况下,高拍仪通过按键方式实现拍照。虚谷号自身没有按键,只是预留有相关接口。若要启动图片采集功能,可以采取使用虚谷号扩展板或通过摄像头采集识别图片或声音中的特殊标志来实现。由于已经采用摄像头来采集图片而且OpenCV视觉处理可以涉及轮廓检测,因此可以划过与被采集场景背景对比强烈的纯色物品,只要虚谷号识别出该物品轮廓面积在一定范围内,就表示开始启动拍照采集。相关关键性程序代码如图

4所示。

3. 工作状态显示

虚谷号兼容Arduino,自带一个可程序控制的LED。从虚谷号开机到准备进入图片采集状态再到图片采集进行中都可以借助该LED的不同频率的亮和灭来表示对应的状态。实现状态显示的代码与借助pinpong库来实现基本一致,亮灯灭灯的次数、亮灯持续时间、灭灯持续时间,可以通

过一个自定义函数实现。例如,ledstatus(led,5,0.2,0.2)表示等待启动;ledstatus(led,3,1,1)表示已采集1张图片;ledstatus(led,2,2,0.5)

```
import cv2 as cv #导入视觉处理模块
count=0
cap = cv.VideoCapture(0) #Opencv3.5 版本适用
cap.set(3,1024) #设置分辨率
cap.set(4,768)
while cap.isOpened():
    flag,frame=cap.read()
    #此处可以考虑对抓拍的图片进行轮廓大小识别,形成触发拍照
    cv.imshow("测试场景", frame)
    k=cv.waitKey(100) & 0xff
    if k==ord('q'):
        break
    elif k==ord('s'):
        count=count+1
        cv.imwrite('upload'+str(count)+'.jpg', frame)
    cap.release()
    cv.destroyAllWindows()
```

图3 图片采集示例代码

```
hsv = cv.cvtColor(frame,cv.COLOR_BGR2HSV)
lower_blue = np.array([100,30,30], dtype=np.uint8) #纯色物品像素 HSB 低值
upper_blue = np.array([130,130,130], dtype=np.uint8) #纯色物品像素 HSB 高值
mask = cv.inRange(hsv, lower_blue, upper_blue)
kernel = np.ones((25,25),np.uint8)
dilation = cv.dilate(mask,kernel,iterations = 1)
contours, hierarchy = cv.findContours(dilation,cv.RETR_TREE,cv.CHAIN_APPROX_SIMPLE)
contoursAreaLow,ContoursAreaUp=35000,60000
for i in range(len(contours)):
    area = cv.contourArea(contours[i])
    if contoursAreaLow <area< ContoursAreaUp and cam_mode==0:
        cam_mode=1
        print("进入拍照模式!")
        break
```

图4 轮廓提取启动图片采集

```
import time #主程序中使用
from pinpong.board import Board,Pin #主程序中使用
Board("xugu").begin() #主程序中使用
led=Pin(Pin.D13,Pin.OUT) #主程序中使用
def ledstatus(led=led,onofftimes=3,ledontime=1,ledofftime=1):
    for i in range(onofftimes):
        led.write_digital(1)
        time.sleep(ledontime)
        led.write_digital(0)
        time.sleep(ledofftime)
```

图5 状态显示示例代码

```
from aip import AipOcr
APP_ID = '25044166'
API_KEY = 'b55fCABS8rZNUg2StcZCPtP'
SECRET_KEY = '6XtG8gTyGHqkKqmnnavyI9Y3gkXZRepr'
#指定构建百度文本 OCR 识别应用必须的 3 个数据
client = AipOcr(APP_ID, API_KEY, SECRET_KEY)
def get_file_content(filePath):    #读取图片的方法
    with open(filePath, 'rb') as fp:
        return fp.read()
image = get_file_content('图片文件路径') #d:/pw/upload01.jpg
#调用通用文字识别（高精度版）
result = client.basicAccurate(image);
lists = result['words_result'] #列表
for lines in lists:
    print(lines['words'])
```

图6 调用百度AI文本识别示例代码

表示图片识别完成,可通过客户端查看识别结果。状态显示相关模块代码使用简单通用适应各类开源硬件的pinpong库完成,如上页图5所示。

4.OCR识别

图片采集后,可以考虑文档高拍仪增强功能实现。从实现途径便捷性与OCR识别效果角度考虑,最终决定还是使用调用百度API的方式来实现OCR识别。访问百度AI开放平台,注册成为百度开发者并创建应用后就能得到相应的AppID、API Key、Secret Key信息。根据图6所示的调用百度AI文本识别示例代码可以完成单张图片中的文字识别。

经过测试,在同等条件下,采用百度API调用方式识别效率比使用开源OCR软件识别效率更高。

5.识别结果呈现

每次文字识别结果可以直接累加存储在单个文本文件中,也可

以格式化为网页文件。针对该存储识别结果的文件可以通过建立Http服务,让任何一台与虚谷迷你电脑联网的客户端查阅。

虚谷自带WIFI而且安装了Python,建立Http服务的代码

很简单而且实用,格式如下:

```
python3 -m http.server
8000
```

该代码含义为指定虚谷迷你电脑当前文件夹为Http服务根目录,指定访问端口为8000。

与虚谷迷你电脑联网的任何一台客户端都可以打开浏览器,输入访问网址(http://虚谷迷你电脑IP:8000/)查阅识别结果。

● 文档高拍仪原型功能改进

经过测试,以虚谷号为核心完成的文档高拍仪原型能够基本完成预设的功能,但在调试过程中还发现存在一些问题,还需要不断迭代完善,如:

- ①摄像头的分辨率不高而且变焦过慢,导致抓拍图像并不清晰;
- ②缺少补光功能导致被拍摄对象本身呈现的亮度不一致,影响后期的文字识别;
- ③适当增加虚谷号的传感控制与状态表示,更有利于操作者理

解操作进程;

④运用OPENCV视觉处理能力有待进一步加强,如调整对比度、锐化、降噪等。

运用虚谷号结合摄像头实现文档高拍仪功能,是一个典型的开源硬件项目,虽然解决的过程有些坎坷,但收获很多。教师在学生的技术基础上,给予学生触手可及的创意方案库,充分调动学生的创意,结合实际场景,让他们不断提出新的需求,不断挑战新的难题,做出实践性很强的解决方案。类似虚谷号的迷你电脑能玩的项目其实非常多,如视觉处理眼睛控制鼠标实现媒体互动、视觉处理无人机禁飞区域判断等,希望有更多的老师带着学生一起探究。e