

Design and Implementation of Intelligent Reading Robot Based on STM32F4

Jiansheng Liu, Zhongjiang Yang, Yinqiu Qin, Chaoxing Zhang, Shuo Li, Yixin Wang

Southwest Petroleum University, Chengdu Sichuan
Email: 384419609@qq.com

Received: Jul. 11th, 2020; accepted: Jul. 23rd, 2020; published: Jul. 30th, 2020

Abstract

Based on the analysis of the reading needs of the special people, such as vision disorders, the existing products and patents of the reading robot, an intelligent reading robot based on STM32F4 is designed and developed, which is mainly composed of motion control system and intelligent reading system. The motion control system uses STM32F4 as the main control chip to realize the three-dimensional space motion and page turning automation of the camera. In the intelligent reading system, OpenCV visual library is used for text image processing and text recognition. At the same time, various modules in Python are used to design real-time interactive reading UI, voice interactive control, text electronic storage, network resource acquisition and other functions. The robot will effectively solve the reading obstacles of the elderly, people with visual impairment or disabled people, and it is suitable for public reading corners in libraries, nursing homes, rehabilitation homes and other places, as well as office places that need a large number of books electronic storage and text scanning. Intelligent reading robot can effectively improve the reading efficiency of the people in need, bring comfort and convenience to the users, and reflect the diversified service concept.

Keywords

Intelligent Reading Robot, STM32F4 MCU, OpenCV, Python, Diversification

基于STM32F4的智能阅读机器人设计与实现

刘建生, 杨忠江, 秦银秋, 张朝兴, 李 硕, 王怡欣

西南石油大学, 四川 成都
Email: 384419609@qq.com

收稿日期: 2020年7月11日; 录用日期: 2020年7月23日; 发布日期: 2020年7月30日

文章引用: 刘建生, 杨忠江, 秦银秋, 张朝兴, 李硕, 王怡欣. 基于 STM32F4 的智能阅读机器人设计与实现[J]. 人工智能与机器人研究, 2020, 9(3): 170-181. DOI: 10.12677/airr.2020.93020

摘要

通过对我国视力障碍等特殊人群的阅读需求, 阅读机器人的现有产品、专利等的分析, 设计研究出了一款基于STM32F4的智能阅读机器人, 该机器人主要由运动控制系统和智能阅读系统组成。运动控制系统采用STM32F4单片机作为主控芯片, 实现摄像头的三维空间运动及翻页自动化。智能阅读系统采用OpenCV视觉库进行文本图像处理以及文字识别, 同时采用Python语言中的各种模块, 设计了与用户端实时交互的阅读UI界面、语音交互控制、文本电子化储存、网络资源获取等功能。该机器人将有效地解决老人、视力障碍人群或残疾人士的阅读障碍, 适用于图书馆、养老院和康复院等地的公共读书角, 以及需要大量图书电子化存储和文本扫描的办公场所。智能阅读机器人能有效地提高了需求人群的阅读效率, 给用户带去了舒适性、便捷性, 体现了多元化的服务理念。

关键词

智能阅读机器人, STM32F4单片机, OpenCV, Python, 多元化

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

根据国家统计局统计数据, 2018年我国60周岁及以上人口已占总人口的比重为17.9%, 达到24,949万人, 到2050年这一比重则将进一步提升为34.1%, 同时中国有接近1700多万名视力障碍残疾人[1]。老龄化问题以及视力障碍问题将进一步阻碍他们与文学世界的交流。

随着经济的快速发展, 人们生活水平不断提高, 人们对生活质量的要求也越来越高, 对智能化产品的接受程度也逐渐提高。但是, 老龄化人口越来越多, 视力障碍的残疾人数量也很多, 他们同样也应该享受阅读的权利。因此, 人们对解决此类阅读障碍的智能产品需求处于一个可观的水平。

2. 行业现状

目前市场上也出现了解决此类问题的概念, 但只是提出了针对性解决理念, 对外也没有公布任何研发进展。市面已有的一些阅读功能产品, 如自动翻书机器人: 只能实现自动翻页, 不能阅读文本[2]; 智能阅读系统: 只能阅读已有的电子文档, 且阅读过程单调、枯燥, 缺乏人性化服务[3] [4]; 文本扫描仪: 具有较为精准扫描文本的功能, 但是需要人为辅助翻书, 人力成本以及体验感较差[5] [6]。综上, 以上产品的功能都较为单一, 不能全面有效地解决老人、视力障碍人群或残疾人士阅读不便的问题。

本文开发的机器人具有较高集成度的模块, 可实现文本自动扫描及翻页、文本图像处理及识别、拟人声阅读、文本电子化存储、语音控制和网络资源获取等功能。环境适应能力强, 能应用多个领域、多种场所[7] [8]。

3. 机器人设计

智能阅读机器人在硬件和软件上均采用模块化设计, 主要由机械结构设计、文本图像处理、运动控制系统、智能阅读系统和智能语音系统构成。

3.1. 机械结构设计

在机械结构设计方面,采用 SolidWorks 进行 3D 建模,实现精准的动力仿真,达到最精准控制水平 [9]。智能阅读机器人的机械结构主要由支撑骨架、书本固定结构、文本扫描翻页结构和辅助翻页装置组成。其实物见所示,三维模型如图 1 所示。

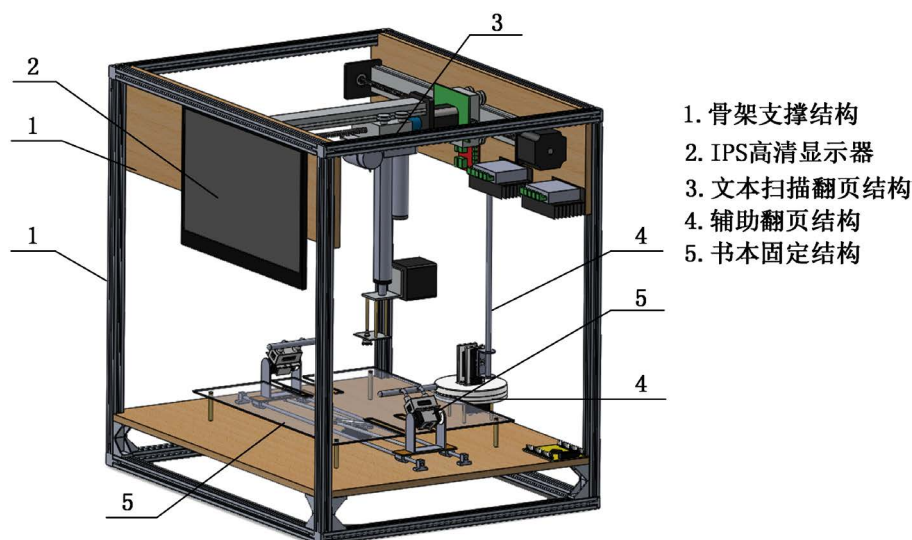


Figure 1. Intelligent reading service robot physical map
图 1. 智能阅读服务机器人实物图

3.1.1. 书本固定结构

书本固定结构由双直线滑块滑杆、数字舵机、曲柄推杆、压书夹和载书平台等组成。固定结构用于防止书本在扫描翻页过程中发生滑动,且能够适应不同大小的书本。

装有滑块的两根滑杆固定在载书平台下方的平台上,每根滑杆两端的滑块用木块连接用于固定数字舵机,使数字舵机能够沿着滑杆左右移动。

压书夹采用 3D 打印技术制成,并固定在数字舵机上。两根滑杆中间为曲柄推杆,可自动控制滑杆两端的压书夹横向移动以固定不同大小的书本。最后将载书平台用铜柱支撑于曲柄推杆上方,载书平台中心设有凹槽,将书脊放于凹槽中可防止书本左右移动。书本固定结构如图 2 所示。

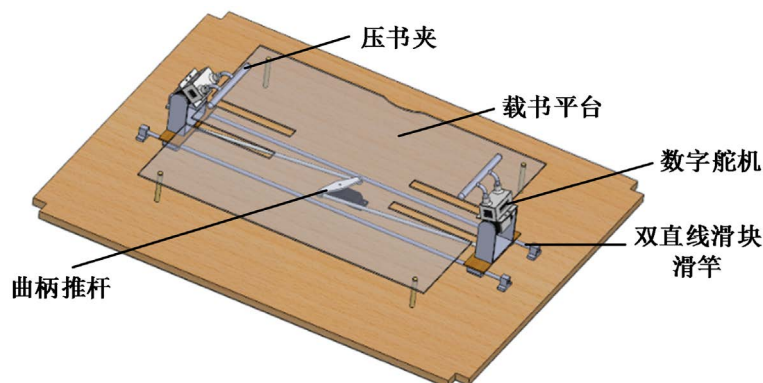


Figure 2. Book fixed structure
图 2. 书本固定结构

3.1.2. 文本扫描翻页结构

文本扫描翻页结构由步进电机、直线导轨滑块滑槽、吸盘真空泵装置、电动伸缩杆、螺纹丝杆、摄像头和继电器等组成。

横向导轨滑槽控制摄像头及吸盘横向运动，纵向导轨滑槽控制其纵向运动，电动伸缩杆控制其垂直运动。通过对步进电机和电动伸缩杆的精确控制可实现摄像头及吸盘的三维空间运动，从而实现文本扫描及自动翻页一体化。文本扫描翻页结构如图 3 所示。

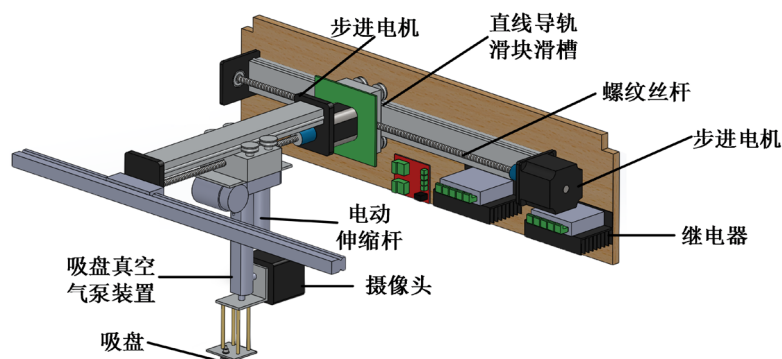


Figure 3. Scan page-turning structure
图 3. 扫描翻页结构

3.1.3. 辅助翻页装置

辅助翻页装置由翻页杆(尼龙杆)、数字舵机和 360°旋转云台组成。旋转云台采用 3D 打印技术制成，其滑槽内放入轴承滚珠以减小摩擦，使翻页杆能够实现旋转运动。当吸盘将书页吸起时，翻页杆运动到吸附的书页下方，防止书页滑落。辅助翻页装置如图 4 所示。

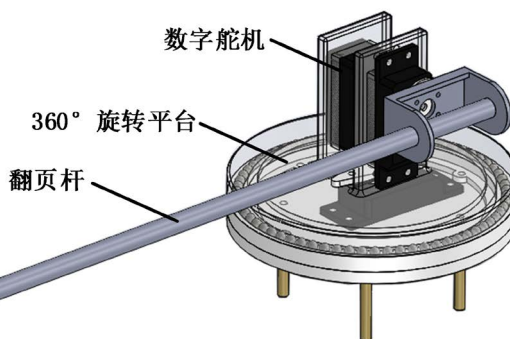


Figure 4. Auxiliary flipping structure
图 4. 辅助翻页结构

3.2. 运动控制系统设计

本作品设计了三种控制模式，分别是自主控制模式、界面控制模式、语音控制模式。整个控制系统框图如图 5 所示。

运动控制系统采用螺纹丝杆构成的三维运动结构；采用 PyQt5 技术实现智能阅读系统 UI 界面的开发与设计；采用 3D 打印技术设计压书夹与辅助翻页云台；使用双曲柄推杆结构实现不同大小书本的固定；使用继电器模块与真空泵实现书本的自动翻页；采用 STM32 控制器和 MBOX 处理器组成的双控制器。运动控制算法代码(部分)如图 6 所示。

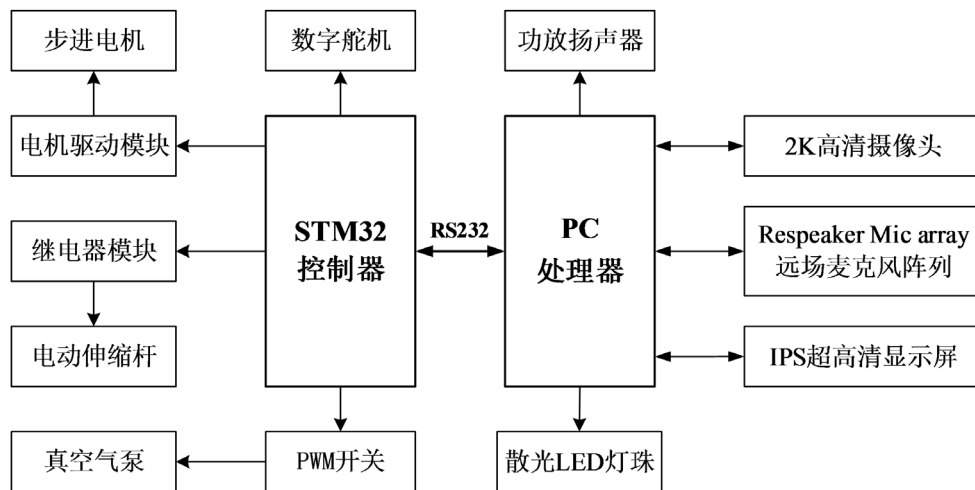


Figure 5. System control block diagram
图 5. 系统控制框图

```
int main(void)
{
    uart_init(115200), delay_init(168);
    LED_Init(), KEY_Init(), BEEP_Init(), EXTIX_Init();
    TIM5_airPump_TimerInit(50 - 1, 8400 - 1);
    TIM12_airPump_Init(5000 - 1, 84 - 1);
    TIM9_setpMotor_X_Init(750 - 1, 84 - 1);
    TIM11_setpMotor_Y_Init(750 - 1, 84 - 1);
    TIM4_Servo_FULR_Init(2500 - 1, 84 - 1);
    TIM3_Servo_B_Init(2500 - 1, 84 - 1);
    digitServo_Callback('F', serAngle_F_begin);
    digitServo_Callback('U', serAngle_U_up);
    digitServo_Callback('L', serAngle_L_begin);
    digitServo_Callback('R', serAngle_R_begin);
    digitServo_Callback('B', serAngle_B_begin);
}
```

Figure 6. Motion control algorithm code (part)
图 6. 运动控制算法代码(部分)

3.3. 软件设计

3.3.1. 文本图像处理

本作品的文本扫描采用 0.3 s 高速自动对焦 2K 高清摄像头，每页文本分批次进行扫描，扫描的原始文本图像如图 7 所示。文本图像处理采用计算机视觉库 OpenCV 的 Python 接口[10] [11]。由于摄像头在安装过程中存在角度偏差，导致采集的原始文本图像在三维空间存在部分文字扭曲变形。因此，使用图像几何变换中的透视变换对原始文本图像进行校正。为解决环境光线对文本图像的影响，使用自适应直方图均衡化调整图像对比度，实现文本图像亮度的均衡化。对处理后的文本图像进行图像形态操作 1，提取文本轮廓，得到文本 ROI 区域，其效果如图 8 和图 9 所示。使用图像形态操作 2 对文本 ROI 区域的每行文字轮廓进行提取，将提取的轮廓与创建的白色背景掩模进行 Poisson 图像融合，再进行自适应直方图均衡化，得到最终的待识别文本，其效果图如图 10 所示[12] [13]。文本图像处理算法框图如图 11 所示。

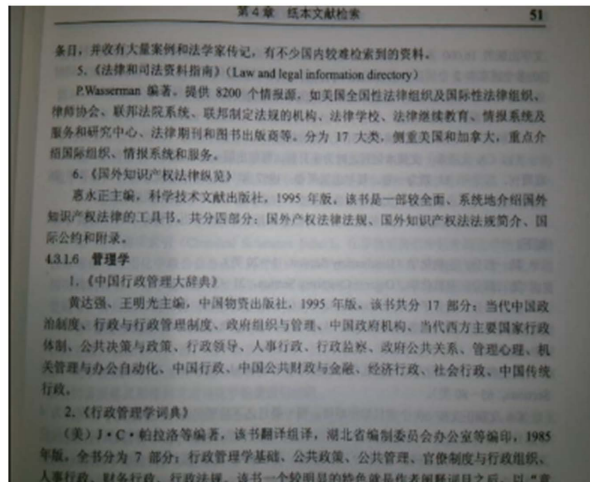


Figure 7. Original text image
图 7. 原始文本图像

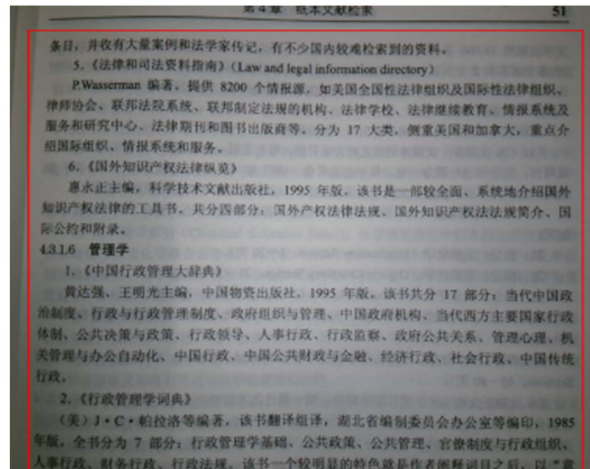


Figure 8. Text outline box
图 8. 文本轮廓框

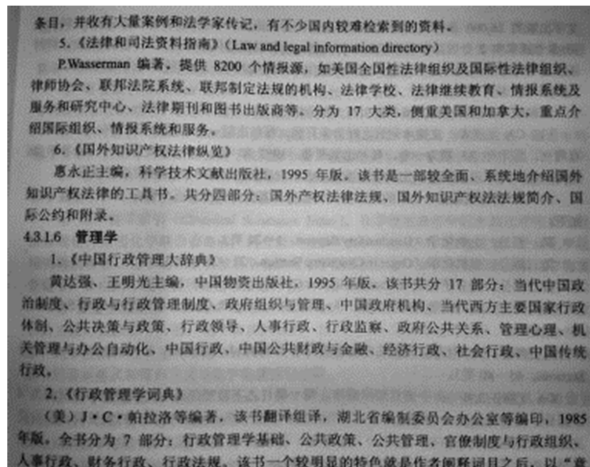


Figure 9. Text ROI area
图 9. 文本 ROI 区域

条目，并收有大量案例和法学家传记，有不少国内较难检索到的资料。

5. 《法律和司法资料指南》(Law and legal information directory)

P.Wasserman 编著。提供 8200 个情报源，如美国全国性法律组织及国际性法律组织、律师协会、联邦法院系统、联邦制定法规的机构、法律学校、法律继续教育、情报系统及服务和研究中心、法律期刊和图书出版商等。分为 17 大类，侧重美国和加拿大，重点介绍国际组织、情报系统和服务。

6. 《国外知识产权法律纵览》

惠永正主编，科学技术文献出版社，1995 年版。该书是一部较全面、系统地介绍国外知识产权法律的工具书。共分四部分：国外产权法律法规、国外知识产权法规简介、国际公约和附录。

4.3.1.6. 管理学

1. 《中国行政管理大辞典》

黄达强、王明光主编，中国物资出版社，1995 年版。该书共分 17 部分：当代中国政治制度、行政与行政管理体制、政府组织与管理、中国政府机构、当代西方主要国家行政体制、公共决策与政策、行政领导、人事行政、行政监察、政府公共关系、管理心理、机关管理与办公自动化、中国行政、中国公共财政与金融、经济行政、社会行政、中国传统行政。

2. 《行政管理学词典》

(美) J·C·帕拉洛等编著，该书翻译组译，湖北省编制委员会办公室等编印，1985 年版。全书分为 7 部分：行政管理学基础、公共政策、公共管理、官僚制度与行政组织、人事行政、财务行政、行政法规。该书一个较明显的特色就是作者阐释词目之后，以“查

Figure 10. To be identified
图 10. 待识别

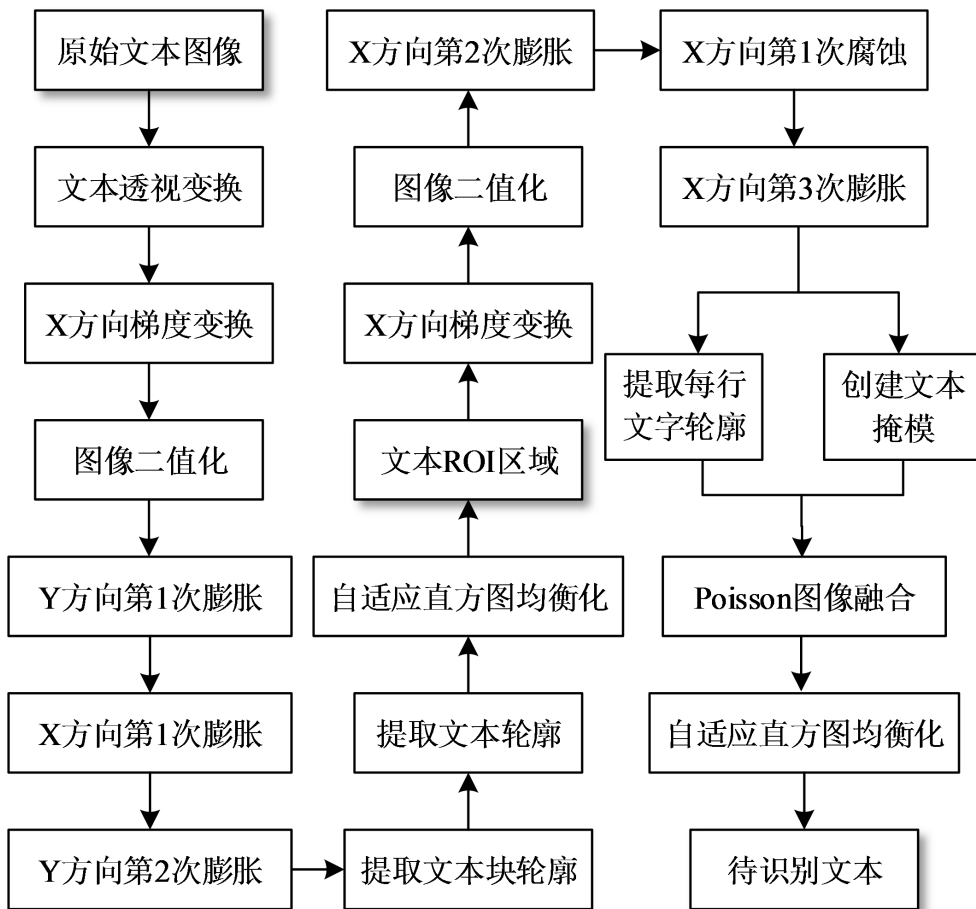


Figure 11. Block diagram of text image processing algorithm
图 11. 文本图像处理算法框图

文本处理程序(部分)如下图 12 所示。

```

import numpy as np
import cv2
from aip import AipOcr
class GetText(object):
    def __init__(self, frame):
        self.srcImg = frame
        self.textImg = frame.copy()
        self.srcGray = cv2.cvtColor(self.srcImg, cv2.COLOR_BGR2GRAY)
        self.eleDila_X = cv2.getStructuringElement(cv2.MORPH_RECT, (20, 3))
        self.eleDila_Y = cv2.getStructuringElement(cv2.MORPH_RECT, (3, 17))
        self.eleEros_X = cv2.getStructuringElement(cv2.MORPH_RECT, (10, 3))
        self.eleEros_Y = cv2.getStructuringElement(cv2.MORPH_RECT, (3, 10))
        self.eleDila_page = cv2.getStructuringElement(cv2.MORPH_RECT, (9, 5))

```

Figure 12. Text processing program (part)

图 12. 文本处理程序(部分)

3.3.2. 智能阅读系统

阅读系统用户端与 STM32 控制器实时交互, 实现用户数据的实时更新, 用户端主要分为智能阅读、智能听书和智能语音三个主要功能模块, 智能阅读模块可以进行页码选择、背景颜色选择和文字大小调节等功能; 智能听书模块主要可以实现播放暂停、音量设置以及设置合成等功能, 其中设置合成主要可以进行人声选择、语音和音调的设置; 智能语音模块可以实现用户的语音识别, 然后反馈给 PC 端。

系统 UI 界面设计使用当前主流界面设计 Qt 的 Python 接口 PyQt5, 其具备开发周期短、实现效果佳和可移植性高等特点[14]。阅读界面主要采用蓝色基调为主, 可以给用户带来视觉上舒适性的体验, 用户可以根据自己的爱好进行相应的系统设置, 从而让阅读者体会到阅读舒适与愉悦, 文本扫描参数设置界面如图 13 所示, 智能阅读系统 UI 界面如图 14 所示。




Figure 13 shows a text scanning parameter setting interface. At the top, it features the title "随风潜入夜, 润物细无声" in a stylized font. Below the title, there are several input fields and buttons:

- "书名:" (Book Name) with an empty text box.
- "起始页数" (Start Page Number) with an empty text box.
- "结束页数" (End Page Number) with an empty text box.
- "总页数" (Total Page Number) with an empty text box.
- "书的大小" (Book Size) with a dropdown menu showing "16K".
- "串口号" (Serial Port) with a dropdown menu showing "COM5".
- A blue "进入" (Enter) button.
- A grey "退出" (Exit) button.

Figure 13. Text scanning parameter setting interface

图 13. 文本扫描参数设置界面



Figure 14. Smart reading system UI interface

图 14. 智能阅读系统 UI 界面

阅读系统算法代码(部分)如下图 15 所示。

```
from PyQt5 import QtCore, QtWidgets, QtGui
import os
import serial
import threading
from GetText_and_img2Text_main import *
import pygame
import Test2Speech
from aip import AipSpeech
import voice_set
from speech_api import speech_recognition
```

Figure 15. Reading system algorithm code

图 15. 阅读系统算法代码

3.3.3. 智能语音系统

语音交互系统设计使用百度在线语音识别和图灵语义解析,系统开始工作时借助 Pyaudio 模块进行录音,然后将录音文件即时上传到 AipSpeech 系统,成功后 PC 端即时接受返回的字符串并匹配字符是否在预设的列表里,若匹配成功,则执行下一步操作,若匹配失败则重新进行录音并按照既定程序工作,最终到语音交互工作结束。同时系统采用 PyGame 的 music 模块,将识别后的文本进行处理,生成文本音频。

综上,采用语音交互系统可以使机器人具有听、思考和说的能力,让整个系统工作起来更加流畅,语音交互系统结构图如图 16 所示。

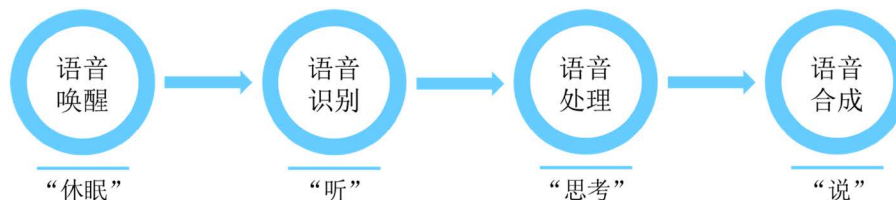


Figure 16. Voice interaction system structure diagram

图 16. 语音交互系统结构图

语音系统代码(部分)如下图 17 所示。

```
import wave
import requests
import time
import base64
from pyaudio import PyAudio, paInt16

framerate = 16000 # 采样率
num_samples = 2000 # 采样点
channels = 1 # 声道
sampwidth = 2 # 采样宽度2bytes
FILEPATH = 'speech.wav'
```

Figure 17. Voice system code (part)

图 17. 语音系统代码(部分)

4. 机器人测试

4.1. 工作过程

2019年10月，完成了该阅读机器人的加工制作与测试，实物图如图18所示。



Figure 18. Robot physical map
图 18. 机器人实物图

对机器人的应用效果进行了测试，机器人的自动翻书结构正在工作，纵向导轨及其末端的吸盘将纸张抓起如图19，纵向导轨将待翻页纸张导引到书本另一端时辅助翻页结构的翻页杆将纸张精准地导引至水平如翻页前图20，一次翻页结束后纵向导轨装置上的高清摄像头正在扫描识别如中图21。

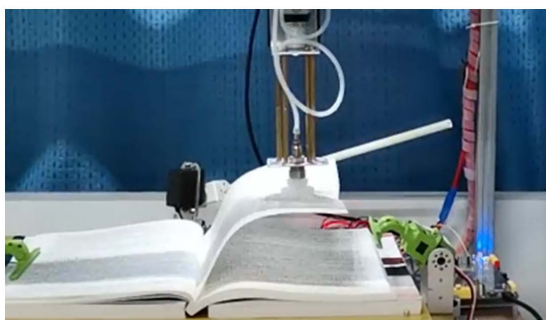


Figure 19. Before page turning
图 19. 翻页前



Figure 20. Page turning in progress
图 20. 翻页中

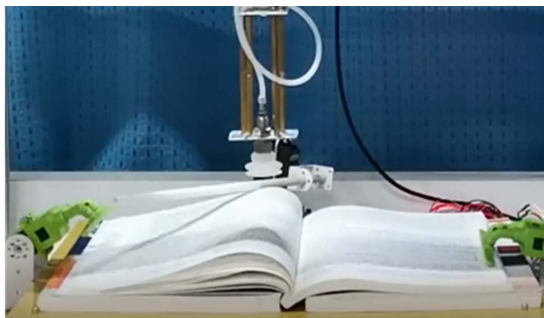


Figure 21. After turning the page
图 21. 翻页后

机器人正在自动识别文本内容并将内容实时反馈到 PC 端，PC 端通过文本图像处理算法计算后，将文本信息显示在高清显示屏的界面上，并向用户语音播报出具体内容。同时，用户也可以采用语音实时控制机器人进行相应动作。

4.2. 测试结果

经过多次测试，智能阅读机器人总体性能良好，通过语音操作或者手动操作选定想要阅读位置后，PC 端可以准确获得位置信息，通过串口将信息传送至主控芯片。主控芯片能够精准地控制电机，使 OpenCV 摄像头运动至指定位置，并拍摄图片信息，通过阅读系统设定的算法，精准地识别了文字信息，最终成功将已识别的文本信息通过串口传送至 PC 端。当机器人遇到外界因素干扰时，系统会自动提醒用户，并自动返回初始状态。机器人测试结果状态如图 22 所示。



Figure 22. Robot test results display
图 22. 机器人测试结果展示

5. 结语

本文为介绍的基于 STM32F4 的智能阅读机器人可以良好地实现文本扫描及翻页自动化、文本图像处理及文字识别、文本电子化存储、拟人声阅读和网络资源获取，同时可实现语音交互控制、UI 界面并行控制等，且运行稳定性好、抗干扰能力强，可应用到多个领域、多种场景，给用户带去舒适的体验感。

参考文献

- [1] 吴伟潮. 浅谈人工智能在中国图书馆行业未来十年的优先主运用[J]. 农业图书情报学刊, 2018, 30(12): 74-77.
- [2] 豆文强, 张禹, 郝晨, 肖敏, 肖建中, 谢太飞, 何小龙, 胡鹏. 自动翻书机器人[P]. 中国专利, CN206588962U. 2017-10-27.

-
- [3] 梁彬, 李静, 高嵩. 一种实现交互式电子阅读的系统、方法和设备[P]. 中国专利, CN106294293B. 2019-10-18.
 - [4] 何孝珍. 一种基于嵌入式的人工智能机器人系统[P]. 中国专利, CN108724159A. 2018-11-02.
 - [5] 唐建波, 李文龙, 赵继印, 马彪, 周春明. 文本自动扫描机[P]. 中国专利, CN202043204U. 2011-11-16.
 - [6] 温悦. 一种文本数据扫描方法和装置[P]. 中国专利, CN108734149A. 2018-11-02.
 - [7] 李艳. 智能阅读服务机器人系统关键技术研究[D]: [硕士学位论文]. 青岛: 中国海洋大学, 2008.
 - [8] 管志伟, 汪明, 周易斌, 王梓洋, 冯雨琪. 一种基于图像相似度计算的阅读机器人[P]. 中国专利, CN107748880A. 2018-03-02.
 - [9] 赵罍, 杨晓晋, 赵楠. SolidWorks 2018 中文版机械设计从入门到精通[M]. 北京: 人民邮电出版社, 2018.
 - [10] Magnus Lie Hetland, 海特兰德, 司维, 等. Python 基础教程[J]. 2014: 2-200.
 - [11] 张平. OpenCV 算法精解[M]. 北京: 电子工业出版社, 2017: 2-398.
 - [12] 邵彩幸, 于越, 段琼, 刘明志. 基于嵌入式与 OpenCV 的目标识别系统研究[J]. 西南民族大学学报(自然科学版), 2018, 44(5): 502-506.
 - [13] 胡文, 马玲玉. 基于 OpenCV 手机拍照快递单文字识别的研究[J]. 哈尔滨商业大学学报(自然科学版), 2015, 31(5): 564-568.
 - [14] 王硕, 孙洋洋. pyqt 5 快速开发与实战[M]. 北京: 电子工业出版社, 2017.