## 《盲人智能辅助系统》开题报告

**小组成员：郭艺君 何妍 刘恩嵩**

1. **产品定义**
2. **选题背景**

盲人是我国人数基数较大的特殊群体，对于如何完善盲人群体的社会服务机制是社会各界热切关注的话题。盲人拥有同样的权利、尊严和价值，同样拥有无穷的智慧、热情与追求，同样是社会物质财富和精神财富的创造者。

基于帮助盲人更好生活，缓解盲人在生活中的诸多不便，我们计划开发一款盲人智能辅助系统，使用嵌入式设备结合云端智能api提供语音读书以及画面描述功能，弥补盲人图书资源的缺口，以及帮助盲人群体具有更独立的生活体验。

1. **选题意义**
2. **帮助盲人在生活中更具有独立性。**一些对于正常人非常简单的事情对于盲人来说可能非常困难，而这可能会给盲人带来沮丧感，如何帮助盲人更具有独立性是面向盲人群体开发的辅助产品的出发点。
3. **解决盲人图书资源的稀缺问题。**在党中央的关怀与重视下，我国对于盲人图书资源已经有一定建设，中国盲文图书馆于2011年6月28日在北京建成开馆，截止2015年，盲文图书已达十万册以上。但盲人图书资源依旧存在很大缺口，且存在分布不均衡等问题。通过将纸质书转换为听书信号是我们项目的一个出发点。
4. **缓解视障人群对于外界的感知能力有限的问题。**视障人群视觉感知能力有限，我们希望通过例如摄像头、超声波这样的传感装置去弥补视障人群一定程度上的感知能力有限问题。
5. **项目目标**
6. **辅助读书功能**

通过摄像头读取面前纸质资源，将图像传入云端智能api，提取出纸质资源中的文字，并通过语音合成技术传回语音信号，通过耳机将文字信息转换成盲人可感知的语音信息。此功能在后续延伸可以用到商品标签识别方面，帮助盲人购物等。

1. **视觉辅助功能**

通过摄像头读取当前图像帧，传入云端智能api（或是本地搭建轻量模型），通过目标检测以及语音合成技术向盲人反馈信息。同时盲人可以通过语音来告知系统想要寻找的物品，系统通过目标识别以及方位确定来告知盲人视野内是否存在目标物品，如果存在即指出其方位。

1. **用户语音控制**

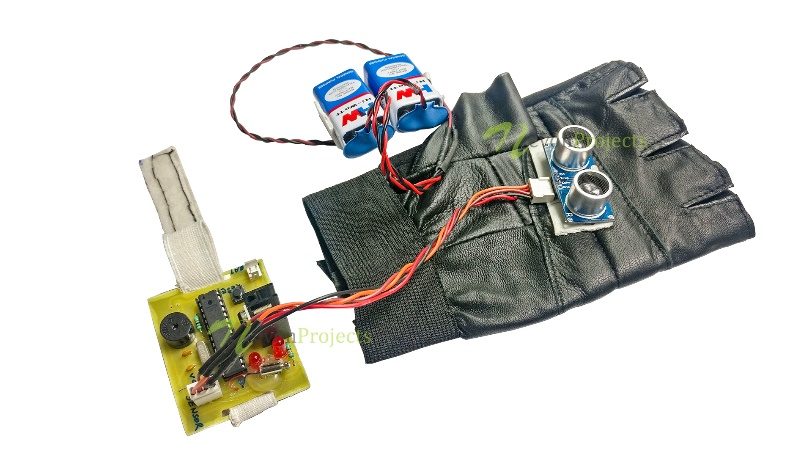
为了向视障人群提供更方便的交互体验，本产品通过语音控制系统执行相关功能，即设计语音指令集来控制系统。后续延伸是提供更多的交互功能。

1. **竞品分析**

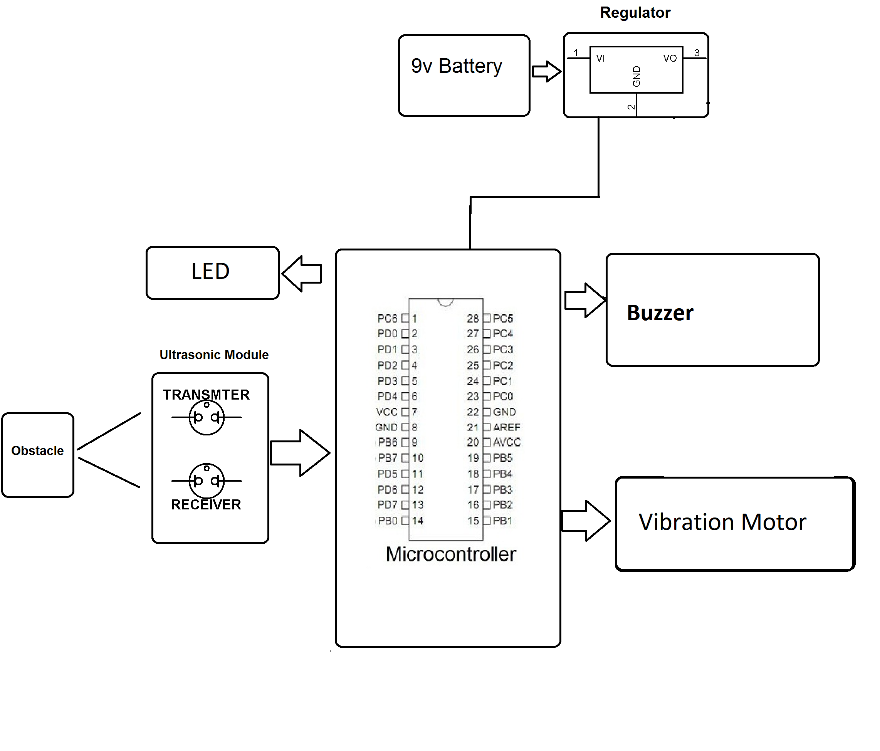
目前针对盲人辅助产品的开发使用到的传感器主要分为三种，第一种是超声波传感器，通过超声波传感器感知前方障碍物以及距离；第二种是摄像头，通过摄像头感知视觉信息弥补视障人群的视觉障碍问题；第三种是深感雷达，通过深感雷达的距离信息感知周围环境，目前使用深感雷达用作无障碍辅助设备的很少，iPhone12首次搭载深感雷达并将其首用于无障碍功能。下面我们将对市场上的现有的盲人辅助系统分为**可穿戴设备**以及**手机上的无障碍功能**。

1. **导盲手套[[1]](#footnote-1)**

该产品通过超声波传感器以及蜂鸣器来进行导盲，离障碍物越近蜂鸣器发生频率越高。功能相对简单单一，基于Atmega-328微处理器实现。



**导盲手套实物图**



**导盲手套原理图**

1. **导盲眼镜**

早期的导盲眼镜通过超声波传感器来实现避障功能。眼镜上装有两只超声换能器和一只耳塞机，换能器能够向前方发射和接收反射回来的超声脉冲波，盲人通过耳机发出的声音变化而感知前方的障碍物。其将[集成电路](https://baike.baidu.com/item/%E9%9B%86%E6%88%90%E7%94%B5%E8%B7%AF)装置安装在眼镜架和眼镜片上，超声装置不断向前发射[超声波](https://baike.baidu.com/item/%E8%B6%85%E5%A3%B0%E6%B3%A2)，当遇到障碍物时即反射回来,经眼镜上的超声波接收装置接收后，通过电子线路的处理，变成可听见的声音,该声音的音调及重复频率是随着障碍物的距离改变而变化，盲人可以根据声音、音调及重复频率来判断前方有无障碍物及障碍物距离。该型眼镜基本原理与导盲手套类似，均是基于超声波传感器，但是这样做的问题在于感知源单一，导航准确率有限，并不能保证安全性，同时也并不具有更进一步的智能功能。[[2]](#footnote-2)

随着深度学习等人工智能技术的迅速发展，目前一些导盲眼镜也具备对于红绿灯识别、斑马线检测等功能，为盲人提供路况信息。[[3]](#footnote-3)[[4]](#footnote-4)

1. **其他可穿戴设备**

瑞士的一家公司开发了Horus可穿戴设备，其设计初衷为“如果改变不了世界各地盲人所处的环境，那就赋予他们生活自如的能力。”Horus设备包含Horus core和Horus eye两部分。Horus core部分只有眼镜盒大小，包含Nvidia Jetson模块、电池等，可以理解为一台迷你计算机。Horus eyes由两颗Full HD摄像头和一对骨传导耳机组成。

视觉障碍者穿戴上Horus设备后，Horus eyes会将其拍摄到的信息传到Horus core中，借助计算机视觉和机器学习技术，经NVIDIA Tegra K1芯片处理后将音频数据回传到Horus eyes的骨传导耳机播放出来，由于音源本身也存在方位，用户可以听出障碍物的大致方向。骨传导耳机则保证了听力障碍者在嘈杂环境下也可正常使用。[[5]](#footnote-5)



**Horus可穿戴设备**

1. **手机的无障碍功能**

手机的无障碍功能是指帮助盲人使用触摸屏手机的一种模式，其改变了常规的触摸屏使用方式，例如通过单击屏幕语音播报所点击的应用来告知盲人所点击的位置，双击打开应用。iPhone与安卓手机目前均已搭载这一功能，iPhone的无障碍模式称为旁白模式，安卓手机的无障碍模式称为Talkback模式。iPhone12在此基础上发展了一个称之为放大镜的功能，其界面与正常相机无异，但是可以语音描述当前摄像头内的画面，是极具人文关怀的一个功能。Google团队也在研发一款叫Lookout的app，其同样是为盲人提供视觉辅助的功能。



**iPhone手机的放大镜功能**

基于以上竞品分析，本产品希望通过嵌入式设备实现一款盲人智能辅助系统，同时结合可穿戴设备的便捷与手机的功能多样性，实现语音听书功能、画面描述功能（类似于iPhone12的放大镜功能）。

1. **系统指标**
   1. **文字识别（包含标点符号等）的准确率大于等于98%**

确保读入信息的准确性。

**测试方法：**通过我们手动采集的书籍内页图片，作为系统测试集，判断是否能达到目标准确率。

* 1. **语音信号的流畅性以及可理解性**

确保盲人使用听书功能的流畅性以及良好的体验，尽可能接近正常的朗读语气与节奏。

**测试方法：**人工测试，判断是否能达到基本标准。

* 1. **目标识别的准确性大于等于98%**

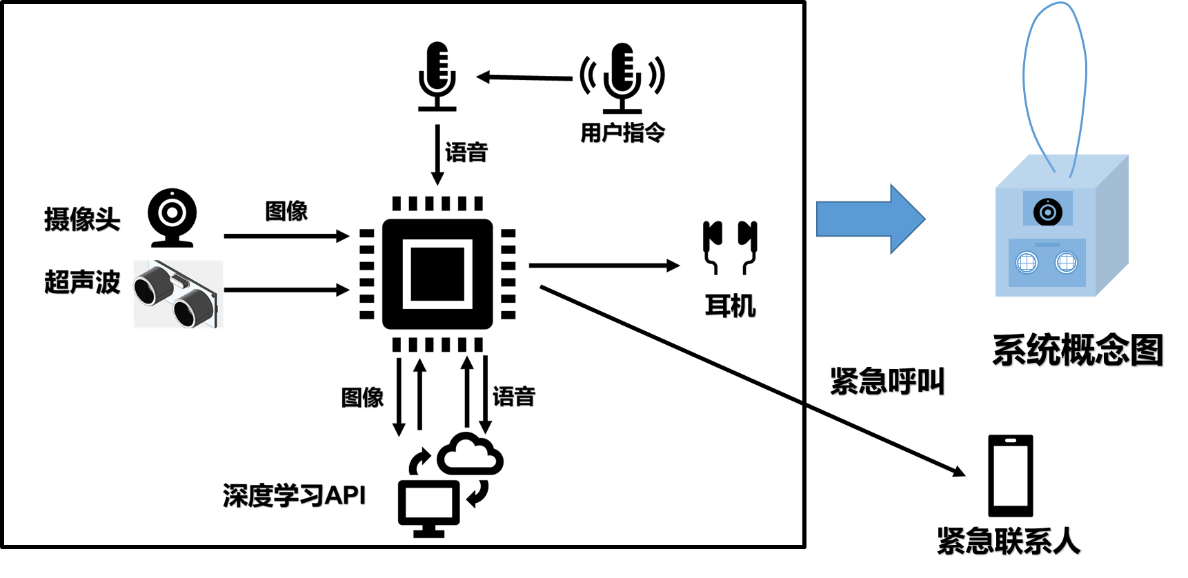
**测试方法：**人工采集现实生活中的场景，送入系统，判断是否能达到目标准确率。

* 1. **目标方位确定及画面描述的准确性（后续）**

**测试方法：**人工采集现实生活中的场景，送入系统，判断对于目标物体的方位识别的是否准确以及画面描述是否可接受。

1. **方案原理**
2. **系统框图**

本产品可大致分为三个模块，分别是云端深度学习模块、传感器信号处理模块与中央控制模块。中央控制模块负责整个系统的启动以及功能选择，传感器信号处理模块将采集的图像、语音、超声波信息处理；云端模块收到图像与语音信号，进行相关的识别再传回中央控制模块，将信息传入文本转语音模块，反馈给用户所需信息。系统框架如下，我们目标是封装成一个可穿戴设备，悬挂在胸前。同时，本产品可以在盲人需要帮助的紧急情况下，通过语音紧急呼叫紧急联系人，这样可以让盲人使用者得到及时的帮助。



**系统框架以及系统概念图**

1. **系统原理**

系统用到的主要技术有超声波反射、图像识别、语音识别以及语音合成。

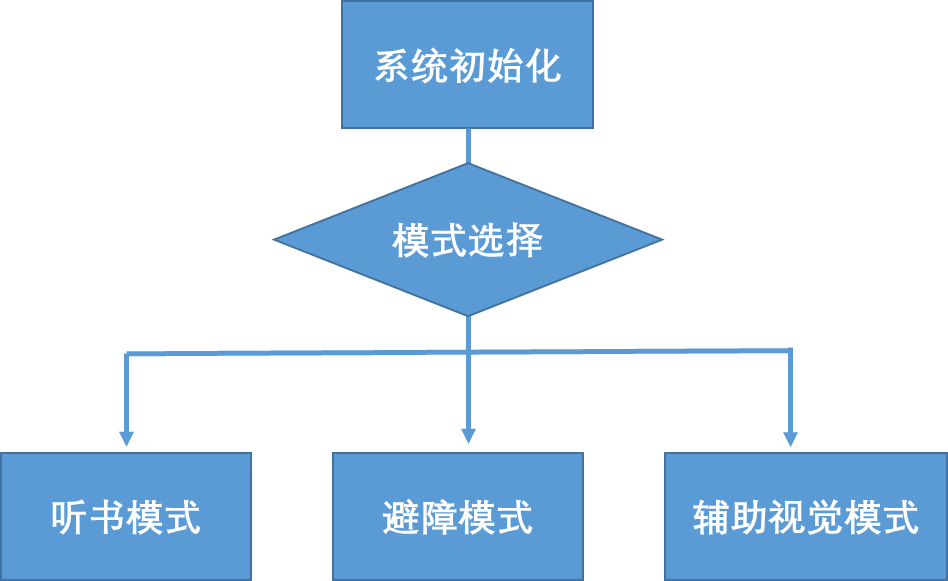
超声波反通过简单的信息处理即可，并无特别的技术难度。

对于图像识别等使用到深度学习的相关技术，由于树莓派算力有限以及功耗问题，无法在本地上完成全部模型的部署，我们想通过调用云端api的方式来进行相关功能的实现。但对于例如描述画面这样有一定实时性要求的功能，可能面临每帧图像都要上传到云端再返回的情境，这就不得不考虑延时的问题，因此，也可以尝试在树莓派本地运行轻量级模型，这样就解决了延时性问题。或者我们降低功能指标，及以更低的刷新频率来换取相对稳定的画面描述功能，当然这就牺牲了一部分的用户实时性体验。

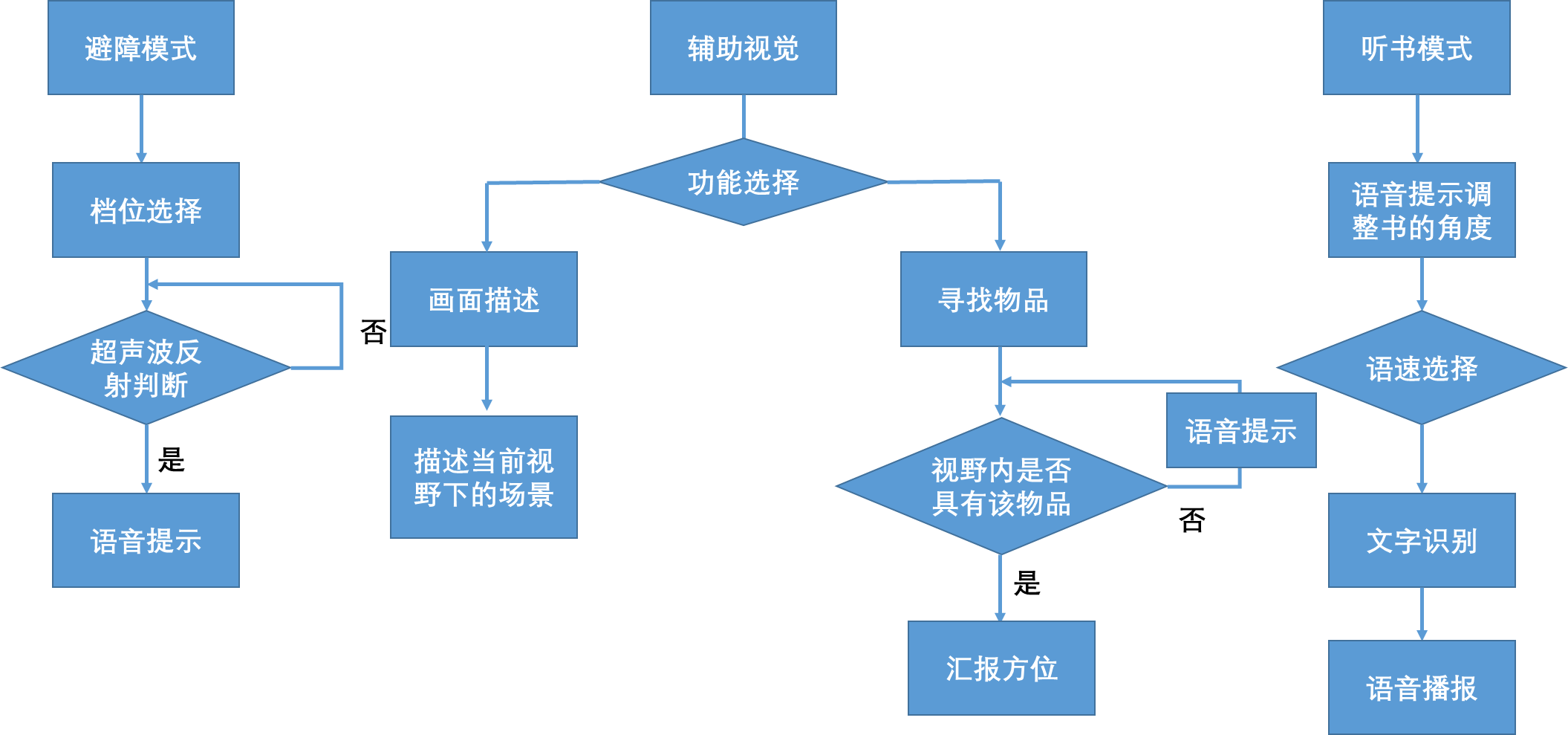
为了实现更好的听书体验，我们可能会在云端获取信息后，对这层信息进行进一步的封装，来实现更专用的功能提升，即更符合阅读图书的节奏。

对于目标的方位识别，这取决于摄像头的选取，对于有深度信息的相机，还可以给出目标的大致距离。对于如何确定方位，可能需要对摄像头进行标定，通过坐标系转化来的到目标的准确定位，或是简单的通过软件层面的处理来给出方位，例如根据目标检测的矩形框的中心在图片上的位置给出大致方位。

我们设计的系统算法流程图大致如下，分为两个层级。



**第一层级流程图**

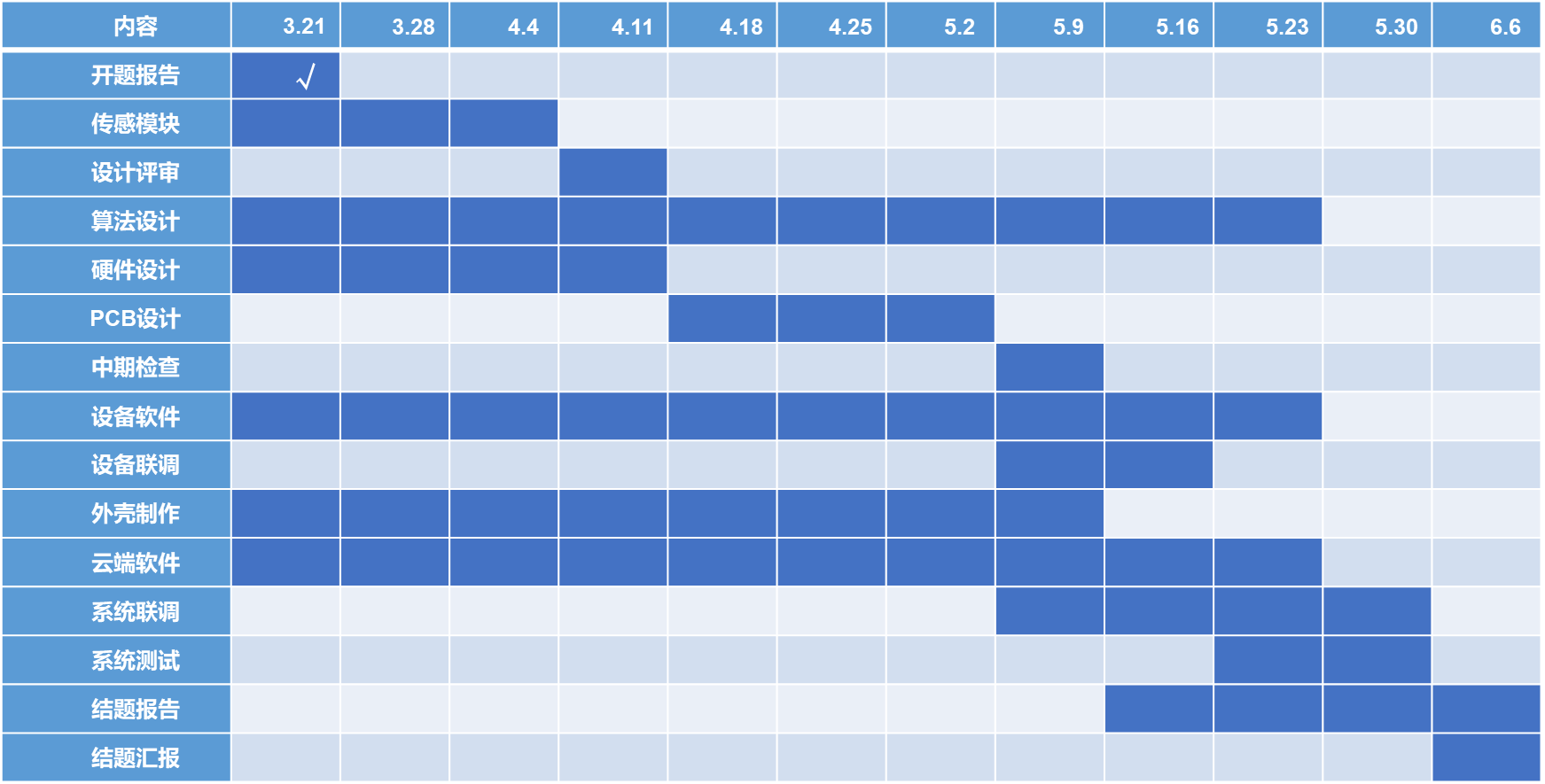
****

**第二层级流程图**

对于系统我们也有进一步的构想，实现针对视障人群在任意环境下的导航功能，我们的初步构想是将目前已有很多研究的同时定位与建图技术（SLAM）结合到系统中，实现系统对周围系统的感知，并实现在感知的基础上，提供定位导航功能。同时在网联基础上，系统可以收集到更多的路端单元信息，实现更安全可靠的功能，为视障人群提供福音。

1. **开发计划**

我们将系统的开发计划分为开题报告、传感模块、设计评审、算法设计、硬件设计、PCB设计、中期检查、设备软件、设备联调、外壳制作、云端软件、系统联调、系统测试、结题报告、结题汇报共15个部分。开发计划甘特图如下所示：



**系统甘特图**

1. **成本预算**

针对该项目的系统方案，我们初步得到如下的成本预算：

|  |  |  |  |
| --- | --- | --- | --- |
| 元件 | 型号 | 产品图片 | 价格 |
| 摄像头 | HBV-1508配zero测试排线 |  | 30.24 |
| 摄像头  备选 带夜视功能 | RPi IR-CUT Camera |  | 91.44 |
| 麦克风 | USBmini microphone |  | 10.00 |
| 超声波传感器 | HC-SR04 |  | 6.00 |
| PCB制版 | 未定 | 未定 | 未定 |
| 外壳 | 未定 | 未定 | 未定 |

**成本预算表格**

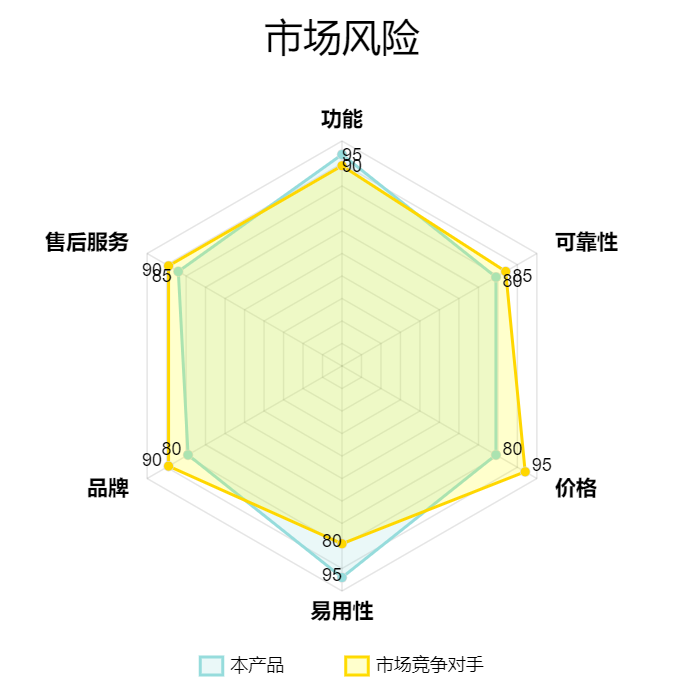
* 1. **项目风险**

1. **工程伦理**

首先，本产品目标以人为本，为盲人的日常生活提供便利，体现了人道主义关怀;其次，本产品特别关注弱势人群的利益，为社会的可持续健康发展提供保障；再次本产品始终保证不损害原则，保证工程项目的质量和安全；最后，本产品节约、节能，充分体现环境友好型原则.

1. **市场风险**

与市场现有竞争对手进行分析对比，在功能方面，我们的产品将具有更全面更多样化的功能；在可靠性方面，由于技术能力尚处于探索发展阶段，可靠性将略低于竞争对手；在价格方面，我们采用低成本的器件设备并极大压低利润，价格将低于竞争对手；在易用性方面，方便快捷、易于使用将是我们产品的一大特征，易用性方面优于市场竞争对手；由于我们尚未具有成熟的品牌与市场，在售后服务、品牌方面，将略逊于市场竞争对手。



**市场风险雷达图**

1. **技术风险**

本产品目前面临三个技术风险：

1. 硬件的设计与集成能否满足小型化需求，能否集成在小型化设备上
2. 在微处理器上调用API或运行程序时，能否满足“盲人智能辅助系统”实时性的需求
3. 如何应对网络状况不佳的状况，以及续航问题。

1. https://nevonprojects.com/third-eye-for-blind-ultrasonic-vibrator-glove/ [↑](#footnote-ref-1)
2. [《中国科技信息》1994年第Z1期 作者：安淑芷](https://baike.baidu.com/reference/7235685/a471J5uaHh7pjaNi-pn1rx0A7VmauK7hwPhTPSbaIAWjrGhZgmJJK9UJPTcnm-xw9HAqzpZElFf_EgvmXrD5JACkyeuaCfyKqhGTd0AKLhA) [↑](#footnote-ref-2)
3. 康亚男, 樊昌熙, 张学毅,等. 双目视觉在智能导盲眼镜中的应用[J]. 电子世界, 2018(9):2. [↑](#footnote-ref-3)
4. 林会祺, 周义涛, 翁名键,等. AI智能导盲眼镜的设计与实现[J]. 信息与电脑, 2021, 33(6):3. [↑](#footnote-ref-4)
5. https://bodyhacks.com/the-horus-wearable-will-allow-the-blind-to-see-again/ [↑](#footnote-ref-5)