勇艺达R150机器人功能分析和设计V0.3

版本号：\_\_\_\_v0.3\_\_\_ \_\_\_

时间：\_\_\_2016.07.15\_\_\_\_ \_

目录

[勇艺达R150机器人功能分析和设计V0.3 1](#_Toc456365496)

[1. 需求分析 4](#_Toc456365497)

[1.1场景布置图 4](#_Toc456365498)

[1.2场景描述 6](#_Toc456365499)

[1.2.1自主导航模式 6](#_Toc456365500)

[1.2.2语音问答模式 6](#_Toc456365501)

[1.2.3特殊导航模式 6](#_Toc456365502)

[1.2.4业务办理模式 6](#_Toc456365503)

[1.2.5门口迎宾模式 7](#_Toc456365504)

[1.2.6 U盘模式 7](#_Toc456365505)

[1.2.7 TTS监控模式 7](#_Toc456365506)

[1.2.8模式之间的关系 7](#_Toc456365507)

[1.3需求分析 8](#_Toc456365508)

[1.3.1地点导航 8](#_Toc456365509)

[1.3.2语音咨询 9](#_Toc456365510)

[1.3.3语音办理业务 10](#_Toc456365511)

[1.3.4迎宾 10](#_Toc456365512)

[1.3.5播放U盘内容 10](#_Toc456365513)

[1.3.6 TTS监控 11](#_Toc456365514)

[1.4界面设计 11](#_Toc456365515)

[1.4.1 APP界面设计 11](#_Toc456365516)

[1.4.2 MTK界面设计 12](#_Toc456365517)

[1.4.3 Intel界面设计 12](#_Toc456365518)

[2.系统架构 15](#_Toc456365519)

[2.2硬件架构 15](#_Toc456365520)

[2.3软件架构： 15](#_Toc456365521)

[3.功能分析 17](#_Toc456365522)

[3.1实时导航： 17](#_Toc456365523)

[3.2语音交互 17](#_Toc456365524)

[3.3环境构建 17](#_Toc456365525)

[3.4后台监控 17](#_Toc456365526)

[4.功能设计 17](#_Toc456365527)

[4.1通信协议 17](#_Toc456365528)

[4.1.1复位 18](#_Toc456365529)

[4.1.2导航命令 19](#_Toc456365530)

[4.1.3开始导航 19](#_Toc456365531)

[4.1.4结束导航 19](#_Toc456365532)

[4.1.5导航错误 19](#_Toc456365533)

[4.1.6行走路径（修改） 19](#_Toc456365534)

[4.1.7 TTS监控开始 19](#_Toc456365535)

[4.1.8 TTS监控结束 19](#_Toc456365536)

[4.1.9 TTS监控文本 20](#_Toc456365537)

[4.1.10遥控信息 20](#_Toc456365538)

[4.1.11障碍信息 20](#_Toc456365539)

[4.2地图构建和标注 20](#_Toc456365540)

[4.2.1登录 20](#_Toc456365541)

[4.2.2账户管理 20](#_Toc456365542)

[4.2.3视频采集及设置 20](#_Toc456365543)

[4.2.4地图生成及显示 21](#_Toc456365544)

[4.3实时导航 21](#_Toc456365545)

[4.4语音对话及语义分析 21](#_Toc456365546)

[4.5 TTS监控： 21](#_Toc456365547)

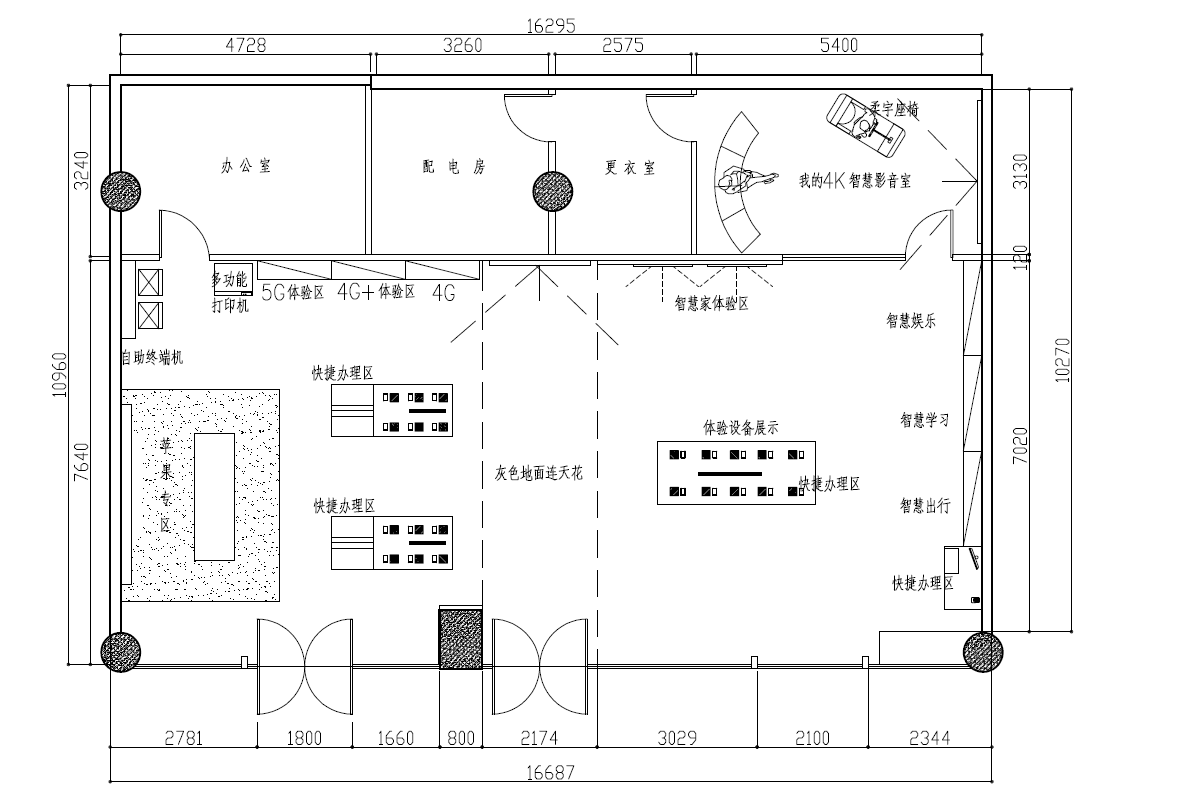
[4.6底盘行动控制和自动避障 21](#_Toc456365548)

[5.开发计划 21](#_Toc456365549)

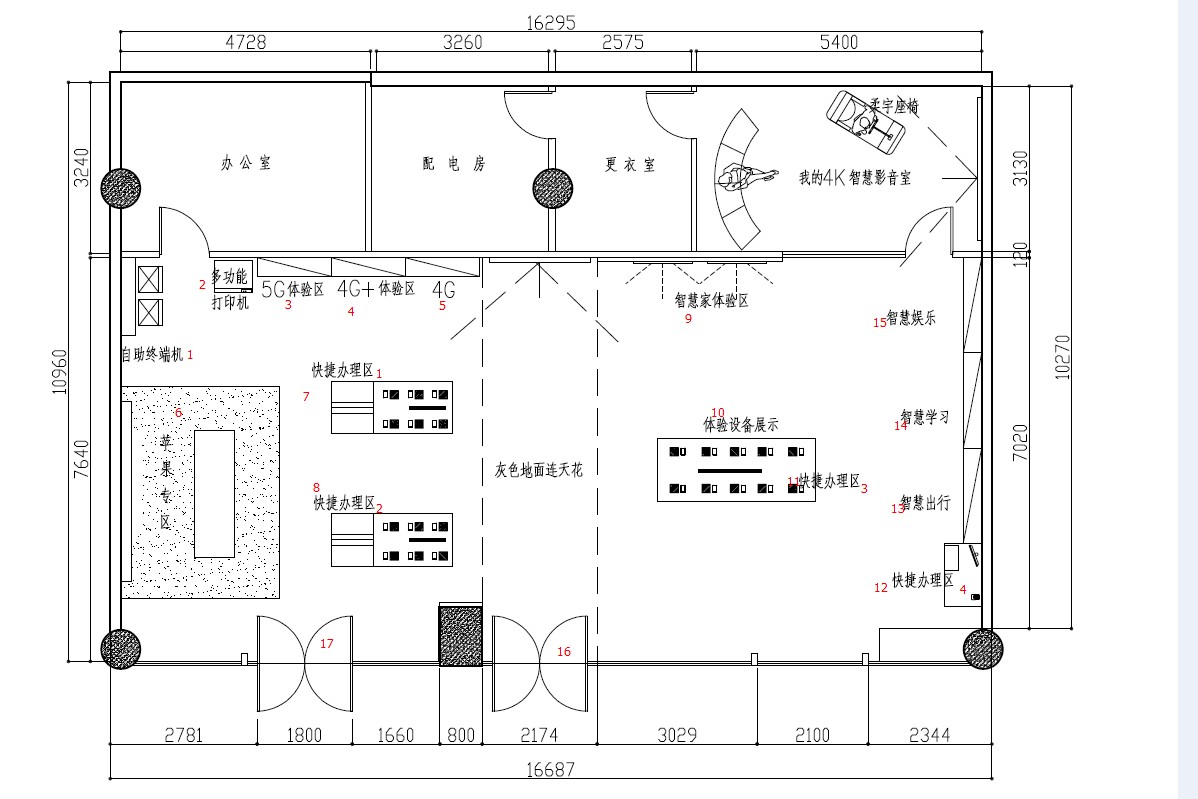
# 需求分析

## 1.1场景布置图

本产品的使用环境为深圳保利广场的移动营业厅，营业厅的室内环境如下图1所示。



本产品服务的营业厅的场景包括：自助终端机、多功能打印机、5G体验区、4G+体验区、4G体验区、快捷办理取1、快捷办理区2、苹果专区、智能家庭体验区、体验设备展示区、智慧娱乐区、智慧学习区、智慧出行去、快捷办理区3、快捷办理区4、大门1和大门2。场景编号如下图2所示：



营业厅的地点编号以及地点标签和说明如下表1所示：

|  |  |  |
| --- | --- | --- |
| **编号** | **地点** | **说明** |
| 1 | 自助终端机 |  |
| 2 | 多功能打印机 |  |
| 3 | 5G体验区 |  |
| 4 | 4G+体验区 |  |
| 5 | 4G体验区 |  |
| 6 | 苹果专区 | 灰色地段为地毯 |
| 7 | 快捷办理区1 | 不需要导航 |
| 8 | 快捷办理区2 | 不需要导航 |
| 9 | 智能家庭体验区 |  |
| 10 | 体验设备展示区 | 不需要导航 |
| 11 | 快捷办理区3 | 不需要导航 |
| 12 | 快捷办理区4 |  |
| 13 | 智慧出行区 |  |
| 14 | 智慧学习区 |  |
| 15 | 智慧娱乐区 |  |
| 16 | 大门1 |  |
| 17 | 大门2 |  |

## 1.2场景描述

### 1.2.1自主导航模式

机器人具备自主行走的能力。用户通过人脸识别或触摸等方式启动语音对话（或者点击win界面上的Icon）之后，机器人能流畅的把客人从前台/门口（或者厅内任意位置）引导到指定的位置。机器人在行走的过程中需要进行绕行避障（或者语音提示并等待路人走开）。在到达指定的地点之后，机器人回到门口/前台的位置。在导航的过程中，路径的选择可以不是最佳。

### 1.2.2语音问答模式

机器人能够对移动业务、产品和服务进行语音问答。用户可以通过摸头等方式启动语音问答。通过摸头等方式启动机器人之后，机器人接收用户的语音输入，并解析结果，之后根据解析结果介绍指定的业务，并在Intel平台上播放相应业务的图片等资料。

### 1.2.3特殊导航模式

机器人还可以自动对展厅进行介绍，机器人初始位置为大厅门口，当接收到自动对展厅进行介绍的指令之后，机器人开始介绍指定的展厅区域，并在介绍完一个区域之后询问用户是否介绍下一个区域。如果是则移动到下一个区域进行介绍，如果是否则监听用户的提问并回答，如果回答否之后识别不到用户的语音输入，机器人会尝试继续询问用户是否进去下一区域。机器人在行走的过程中还会边走边介绍其他内容。直到全部介绍完之后，机器人将停留在原地等待用户的命令，当长时间没有用户与机器人对话之后，机器人将回到门口位置。

### 1.2.4业务办理模式

用户点击Intel平台界面上的Icon之后，MTK对用户点击的业务类型进行语音回答，并且打开办理业务的网页界面，以供用户完成业务办理。

### 1.2.5门口迎宾模式

机器人位于门口位置，当有用户在机器人的正前方时，机器人通过人脸检测检测到用户之后将播放欢迎语句，之后开始监听用户的语音输入并回答用户的问题。当用户没有回答之后，机器人重新开启人脸检测，继续等待下一个用户的到来。

### 1.2.6 U盘模式

用户可以通过U盘将需要播放的资源拷贝到Intel平台上，通过设置资源路径，程序可以对选中的资源进行多种模式的播放。

### 1.2.7 TTS监控模式

TTS监控模式通过两种方式实现：一是直接的文本转换，二是通过变声。

直接文本转换的流程为：用户通过手机APP绑定机器人，然后通过手机APP打开监控模式。此时用户可以遥控机器人的行走方向，也可以对机器人周围的视频和音频信息进行实时监控。当听过机器人周围的人说话时，用户可以通过点击屏幕的按钮将用户的声音转换为文本并传送到机器人端播放出来。

而通过变声则不同，当监控建立之后，用户的声音将实时的传送到机器人端，机器人端将用户的声音变声之后播放出来。

### 1.2.8模式之间的关系

系统默认为自主导航模式。

U盘模式和TTS监控模式为独立存在的模式，当进入这两种模式之后，其他模式将不能正常工作。当用户有打断意图时，机器人将播放提示声音提示当前不能进行别的任务。

自主导航模式、语音问答模式、业务办理模式可以同时存在。模式的选择根据语义解析的结果进行切换。

特殊导航模式、语音问答模式可以同时存在，但可以被TTS监控模式打断。

门口迎宾模式、语音问答模式、业务办理模式可以同时存在。

模式的优先关系为：TTS监控模式>特殊导航模式>U盘模式>自主导航模式>语音问答模式=业务办理模式。高优先级意味着可以打断低优先级的任务，并切换到高优先级模式下工作。

## 1.3需求分析

### 1.3.1地点导航

机器人的初始位置为前台/门口，当检测到有人在机器人的正前方并且停留一定时间之后，机器人将通过语音与客户打招呼，之后询问客户需要办理什么业务，在得到用户需要办理的业务类型之后，机器人开始移动，并将用户引领到目标地点并进行语音提示，在完成这一系列动作之后之后，机器人将回到前台/门口的位置。机器人需要导航的位置包括：自助终端机、多功能打印机、5G体验区、4G+体验区、4G体验区、快捷办理取1、快捷办理区2、苹果专区、智能家庭体验区、体验设备展示区、智慧娱乐区、智慧学习区、智慧出行去、快捷办理区3、大门1和大门2。本需求的流程图如下图2所示：



### 1.3.2语音咨询

对于产品和服务的介绍通过语音播放的方式呈现（是否需要图片展示）。机器人在门口或者其他位置固定，随机通过语音播放一些产品的介绍。当用户通过对话的触发产品或服务的介绍时，机器人将播放产品或服务的介绍语音。本功能业务流程图如下图3所示：



### 1.3.3语音办理业务

Intel平台展示业务办理界面，当用户点击界面上的图标时，机器人开始为用户介绍业务办理的流程并且在Intel平台上打开相应的网页。

### 1.3.4迎宾

当开始迎宾模式之后，机器人将关闭自动导航定位功能。机器人将固定在一个地点，当通过人脸识别检测到客户之后将播放欢迎语句并与用户进行语音对话。

### 1.3.5播放U盘内容

当用户插入U盘之后，可以将U盘指定文件夹内容拷贝到Intel平台上。用户可以通过程序的设置功能查看分类显示的资源列表（视频和图片）；用户可以选中需要播放的多个视频或者多个图片，并选择播放模式，之后程序将对多个视频或者图片进行循环播放。

### 1.3.6 TTS监控

用户可以通过APP脸上机器人，机器人进入监控模式，当客户与机器人进行对话时，客户的音频将发送到APP端，用户听到之后通过语音转文字功能，将文字发送到机器人段，机器人端通过语音合成功能播放出文字。这样完成了TTS监控功能。同时，TTS监控模式下还能对机器人进行视频监控和遥控。遥控可以控制机器人的行动方向。本功能如下图4所示：



## 1.4界面设计

### 1.4.1 APP界面设计

APP端的功能主要为音视频监控、TTS监控、远程遥控。APP界面如下图5所示：



### 1.4.2 MTK界面设计

MTK平台界面始终显示为表情界面（wifi设置），如下图6所示。



### 1.4.3 Intel界面设计

Windows端的主界面包括业务区、菜单区。菜单区包含菜单、模式、帮助等菜单：菜单包含退出选项；模式包括迎宾模式、业务导航模式、广告模式。帮助包含了系统的部分功能，比如地图建立、地点标注、系统设置等。

图7

#### 1.4.3.1地图构建界面设计

地图的构建以离线的形式进行，界面如下所示：

（1）用户登录界面视：



1. 账户管理：



（3）视频采集及设置：



1. 地图生成及显示：



# 2.系统架构

## 2.2硬件架构

本系统硬件架构分为电源部分、MTK8735平台部分、Intel平台部分、STM32部分。

图8

系统采用统一电源供电，MTK和Intel平台采用统一的电源管理。MTK平台包括MIC阵列、触摸屏、喇叭、摄像头、蓝牙、3G/4G模块、WIFI模块等。Intel平台集成了触摸屏、RealSense等模块。STM32平台控制底盘各个传感器模块，包括电机、超声波等。MTK平台、Intel平台和STM32平台之间通过串口进行通信。（需要给出详细的）

## 2.3软件架构：

系统分为：MTK模块、Intel模块、硬件控制模块、服务器模块、APP模块。如下图所示：



图9

其中，Server模块包括用户统计模块、语义模块、后台管理模块。用户统计模块用于统计系统的运行状态，包括语音对话、系统错误信息、系统升级等。语义模块负责对用户输入的语句进行解析。后台管理模块负责系统的配置，如系统升级等。

APP模块包括视频监控、音频监控、遥控。视频监控是对机器人的的摄像头视频信息的监控。音频监控是对机器人的麦克风音频信息进行监控。遥控是对机器人进行远程遥控，控制机器人的运动方向。

Hardware模块是硬件控制模块，负责对硬件进行控制和反馈。其中底盘模块可以同时接收来自Intel和MTK平台的控制信息。

MTK平台负责表情显示和语音处理功能，分为通讯模块、业务模块、传感器模块。通讯模块功能包括网络通信、串口通信。业务模块提供基础的业务功能实现，包括：语音转文本、文本转语音、语义理解、生源定位、TTS监控、视频监控、语音监控等等。传感器模块是MTK平台上集成的各个传感器组件，负责对机器人周边信息的采集和传输。

Intel平台负责算法的实现以及系统UI的实现。其中通讯模块负责网络通讯和串口通讯。业务模块包括厅店介绍、自助迎宾、优惠介绍、自动升级、基本娱乐、厅内导购、业务推荐、客户导航以及其他业务。算法模块负责算法部分的实现。传感机器模块集成了Intel模块的传感器，负责对机器人周边信息的采集以及对机器人的控制。UI显示模块负责UI显示，提供用户交互界面。

# 3.功能分析

本产品（R150）的使用环境为中国移动营业厅，主要功能为：用户业务的实时导航、语音交互、环境构建、后台监控。

## 3.1实时导航：

用户业务实时导航是指，在机器人构建好营业厅的环境之后，当用户与机器人进行语音时，机器人通过识别业务类型，获得用户需要到达的目的地点，之后将带领用户到达目的地，并且在运行的过程中进行语音提示和避障。

## 3.2语音交互

用户语音交互指的是与语音相关的交互，用户通过语音与机器人进行交互，机器人根据用户的语音音输入做出相应的反馈，按照业务相关分为与业务相关的语音交互和与业务无关的语音交互。

与业务有关的语音交互包括：导航类的业务和咨询类的业务。导航类的业务需要进行实时导航，将用户引导到业务办理的目的地，并进行语音提示；咨询类的业务只需要进行语音交互即可。

## 3.3环境构建

环境搭建指的是系统的初始化，包括对营业厅的地图构建、地点标定、以及其他一些的系统配置。

## 3.4后台监控

后台监控包括视频监控、语音监控、TTS监控、遥控等。视频监控是对机器人前方的视频信息进行监控。语音监控是对用户的语音进行监控。TTS监控是取消机器人的自动回答，通过将后台人为控制，将语音转化为文字，再通过R150机器人通过语音合成功能播放出来。遥控是对机器人的远程控制。

# 4.功能设计

## 4.1通信协议

系统采用JSON格式作为本系统内部模块之间的通信协议。系统内部各模块之间的通讯示意图如下所示：



图10

各协议如下表2所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 编号 | 说明 | 流向 | 备注 |
| 10 | 复位命令 |  | 预留 |
| 300 | 导航命令 | MTK🡪Intel |  |
| 301 | 开始导航 | Intel🡪MTK |  |
| 310 | 结束导航 | Intel🡪MTK |  |
| 302 | 导航错误 | Intel🡪MTK |  |
| 303 | 行走路径 | Intel🡪STM32 |  |
| 401 | TTS监控开始 | MTK🡪Intel | 关闭人脸识别 |
| 410 | TTS监控结束 | MTK🡪Intel | 开启人脸识别 |
| 420 | TTS监控的文本 | APP🡪MTK | MTK播放文本 |
| 501 | 遥控信息 | APP🡪MTK🡪STM32 | 机器人的行动信息 |
| 601 | 前方有障碍物 | STM32🡪Intel |  |
|  |  |  |  |

表2

协议格式为：{“cmd”:1,”content”:{“speed”:”100”,”oritention”:”1”}}，其中，”cmd“为命令的编号，”content“为命令的具体内容。

### 4.1.1复位

各模块接收到复位命令之后将停止正在进行的业务，回到初始化状态（表情显示平常表情、WIN界面回到页面显示状态、语音中止播放）。命令格式为：{“cmd”:1,”content”:”null”}。

### 4.1.2导航命令

由MTK平台发送到Intel平台。命令格式为：{“cmd”:300,”content”:{“des”:”1”}}。”des”指目标地点。

### 4.1.3开始导航

由Intel平台发送到MTK平台。当算法模块规划好路径之后开始导航时发出。命令格式为：{“cmd”:301,”content”:{“des”:”1”}}。”des”指目标地点。

### 4.1.4结束导航

由Intel平台发送到MTK平台。当到达目的地点终止导航时发出。命令格式为：{“cmd”:310,”content”:{“des”:”1”}}。”des”指目标地点。

### 4.1.5导航错误

由Intel平台发送到MTK平台。当导航出错时发出。命令格式为：{“cmd”:302,”content”:{“des”:”1”}}。”des”指目标地点。

### 4.1.6行走路径（修改）

由Intel平台发送到MTK平台。算法模块规划好路径之后需要驱动底盘运动到目标地点。命令格式为：{“cmd”:303,”content”:{“lu”:1, “ld”:1, “ru”:1, “rd”:1, “lut”:1, “ldt”:1, “rut”:1, “rdt”:1 }}。其中：lu表示左(left)上(up)轮的运行速度，ld表示左下轮的运行速度，ru表示右上轮的运行速度，rd表示右下轮的运行速度，lut表示左上轮运行时间，ldt表示左下轮运行时间，rut表示右上轮运行时间，rdt表示右下轮运行时间。其中时间的单位为ms,速度的单位为m/s。

### 4.1.7 TTS监控开始

由MTK平台发送到Intel平台。命令格式为：{“cmd”:401,”content”:”null”}。

### 4.1.8 TTS监控结束

由MTK平台发送到Intel平台。命令格式为：{“cmd”:410,”content”:”null”}。

### 4.1.9 TTS监控文本

由APP发往MTK平台。命令格式为：{“cmd”:420,”content”:{“text”:”我的名字叫小勇”}}。

### 4.1.10遥控信息

由APP发出，MTK平台转发到Intel平台。命令格式为：{“cmd”:501,”content”:{“lu”:1, “ld”:1, “ru”:1, “rd”:1, “lut”:1, “ldt”:1, “rut”:1, “rdt”:1 }}。其中：lu表示左(left)上(up)轮的运行速度，ld表示左下轮的运行速度，ru表示右上轮的运行速度，rd表示右下轮的运行速度，lut表示左上轮运行时间，ldt表示左下轮运行时间，rut表示右上轮运行时间，rdt表示右下轮运行时间。其中时间的单位为ms,速度的单位为m/s。

### 4.1.11障碍信息

由STM32发送到Intel平台。命令格式为：{“cmd”:601,”content”:{“c1”:1,”c2”:1,”c3”:1,”c4”:1,”c5”:1,”c6”:1,”c7”:1,”c8”:1}}。其中c1🡪c8为各个超声波接收到的距离信息。

## 4.2地图构建和标注

地图的构建和标注是系统的基本功能，是其他功能实现的前置功能。

### 4.2.1登录

使用地图的构建和标注功能的前置条件是需要用户进行登录。当登录成功之后，系统进入到地图构建和标注的主界面；如果失败则提示错误信息，并停留在登录界面。

### 4.2.2账户管理

对用户的账户密码进行维护，可以对用户账号密码进行修改。

### 4.2.3视频采集及设置

### 4.2.4地图生成及显示

## 4.3实时导航

实时导航过程不接受其他的功能切换，必须执行完之后才能去进行其他的功能，并且导航的过程中具有避障功能。

## 4.4语音对话及语义分析

语音对话分为导航型和非导航型。导航型需要机器人在判断出目标地点之后导航到相关的地点，而非导航型业务只需要通过语音回答。

## 4.5 TTS监控：

TTS监控需要APP通过网络与MTK平台连接，APP可以监控MTK平台的视频和音频信息，APP通过语音转文本将用户的语音转换为文本并转发到MTK平台，MTK平台接收到文本之后通过语音合成播放文本。在监控的同时，APP还可以控制机器人的行动方向。

## 4.6底盘行动控制和自动避障

在接收到上层的控制命令之后，底盘可以进行移动。在运动的过程中还可以将超声波信息上传到上层应用中。

# 5.开发计划

通过前面章节的分析和设计，可以得到本系统的功能划分表如下所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **模块** | **功能划分** | **功能说明** | | **责任人** | **开发计划** |
| APP | 视频/音频监控 | 对机器人面前的视频和音频信息进行监控 | | 章工 |  |
| 行动控制 | 控制机器人的行动 | | 章工 |  |
| 对话监控（文本和变声） | 通过语音转文本（或者变声）并发送到机器人的MTK端，完成与用户的交互。 | | 章工 |  |
| MTK | 主服务 | | 对机器人的语音对话进行控制，通过与用户之间的对话控制机器人做出相应的回应。 | 章工 |  |
| 表情图片替换 | | 需要将图片换成中国移动指定的图片 | ready |
| 业务逻辑控制模块 | | 保证系统的正常运行 | 肖 |
|  | |  |  |
|  | |  |  |
| Intel | UI | | 包括主程序UI以及地图构建UI | 杜工 |  |
| 算法模块（定位导航、人脸识别） | | 实现系统所需要的各个算法，并且提供接口调用。 | 明明  汪工 |
| 业务逻辑控制模块 | | 保证系统的正常运行 | 杜工 |
| 地图构建 | |  | 李湘 |
|  | |  |  |
| Base | 超声波避障 | |  | 黄工  向颖 |  |
| 电机控制 | |  |
| 电源控制 | |  |
|  | |  |
|  | |  |
| 协议 |  | | | 肖 |  |
|  |  | | |  |  |
|  |  | | |  |  |