

第五届  
全国大学生集成电路创新创业大赛

参赛杯赛\*: 平头哥杯

作品名称\*: 基于玄铁 E902 的智能搜救机器人

队伍编号\*: CICC3235

团队名称\*:

## 摘要

救援现场的复杂性和潜在危险因素给救援人员带来极大安全隐患,因此很多国家都对灾害救援场景下的救援机器人开展了大量研究。无人车、无人机等救援机器人凭借着高性能、高灵活性、高精度、在危险区域执行任务不造成人员伤亡等特点,被广泛应用于各种抢险救灾现场、侦查打击等现场、以及日常的危险监视领域。

由于很多灾后现场的供电设施和通信基站处于瘫痪状态,导致目前通用 4G 网络在灾后地区处于微弱信号或无法使用的情况。这种通信不稳定以及动态的环境下进行,会进一步导致救援机器人在不确定性通信条件下进行作业,从而会大大影响灾后救援作业的效率。因此提高机器人系统在通信不稳定的情况下能正常与主控台进行通信显得尤为重要。本团队设计一款基于应急通信情况下的救援机器人,以解决应急情况下通信方式与救援需求不匹配的问题。

本团队拟设计的应急通信救援机器人可选择两种通信模式,一种是正常的 4G 网络通信,该通信模式距离无限制。一种是在无法使用 4G 网络的时候,可以基于 WIFI Halow 通信方案进行 1 公里以上的自由稳定通信,不被基站所限制,可以满足所有灾后现场的特殊通信情况。除此之外,本团队设计的应急通信救援机器人可以在灾后特殊环境搜救被困人群,还可以将物资精准运送到指定地区,进行物资的分发。本救援机器人还具有地图定位功能,能在第一时间将被困人员的具体位置发送给救援人员。同时还可以将采集到的音视频数据和一些传感器数据如温湿度、气体检测数据通过 WIFI 或者 4G 传输给控制端,上位机也可以远程遥控机器人的运动与现场进行音视频通话。本设计采用基于 RISC-V 架构的 E902 处理器来实现的,通过在 E902 上挂载加速 IP 实现音视频数据的实时处理、传感器数据采集并实现了上位机控制端与救援机器人之间的高效远程数据传输。

**关键词:** 音视频处理; Wi-Fi Halow; RISC-V; 实时传输

## **Abstract**

The complexity and potential danger of rescue sites pose a great safety risk to rescuers, so many countries have conducted a lot of research on rescue robots in disaster rescue scenarios. Unmanned vehicles, drones and other rescue robots are widely used in various rescue and disaster relief sites, detection and combat sites, and daily hazard surveillance areas because of their high performance, high flexibility, high accuracy, and the ability to perform tasks in hazardous areas without causing casualties.

As many post-disaster site power supply facilities and communication base stations are down, resulting in the current general-purpose 4G network being in a weak signal or unusable situation in post-disaster areas. This communication instability as well as the dynamic environment in which it takes place can further lead to rescue robots operating under uncertain communication conditions, which can greatly affect the efficiency of post-disaster rescue operations. Therefore it is particularly important to improve the robot system to be able to communicate normally with the main console in the case of communication instability. This team designs a rescue robot based on emergency communication situations to solve the problem of mismatch between communication methods and rescue needs in emergency situations.

The emergency communication and rescue robot to be designed by this team can choose two communication modes, one is the normal 4G network communication, the distance of this communication mode is not limited. One is when 4G network is not available, it can communicate freely and stably based on WIFI Halow communication scheme for more than 1 km, which is not limited by base stations and can meet all special communication situations in post-disaster sites. In addition, the emergency communication and rescue robot designed by our team can search and rescue trapped people in special post-disaster environments, and can also accurately transport supplies to designated areas for distribution of supplies. This rescue robot also has map positioning function, which can send the specific location of trapped people to rescuers in the first time. It can also transmit the collected audio and video data and some sensor data such as temperature and humidity, gas detection data to the console through WIFI or 4G, and the host computer can also remotely control the robot's movement and make

audio and video calls to the scene. The design is based on the RISC-V architecture of the E902 processor to achieve the real-time processing of audio and video data, sensor data acquisition and efficient remote data transmission between the host console and the rescue robot by mounting accelerated IP on the E902.

**Key words:** Audio-video processing; Wi-Fi Halow; RISC-V; real-time transmission

# 目 录

第 1 章 作品概述 .....	1
1.1 作品意义 .....	1
1.2 作品简介 .....	2
第 2 章 作品难点与创新 .....	4
2.1 作品难点 .....	4
2.2 关键技术研究 .....	4
2.2.1 实时监控信息通信技术研究 .....	4
2.2.2 音视频实时同传技术研究 .....	4
2.3 作品创新 .....	5
2.3.1 RISC-V SOC 设计 .....	5
2.3.2 无线高清音视频实时同传技术 .....	5
2.3.3 高可靠远程遥控机器人传技术 .....	5
第 3 章 硬件设计与流程 .....	6
3.1 RISC-V SOC 设计 .....	6
3.2 G.711 音频编码加速 IP .....	7
3.2.1 G.711 编码原理分析 .....	7
3.2 火焰识别算法加速 IP .....	9
3.2.1 识别算法原理分析 .....	9
3.2.2 识别算法实现 .....	10
第 4 章 软件设计与流程 .....	14
4.1 E902 CDK 工程实现 .....	14
4.2 OK6818 设备平台程序实现 .....	14
4.3 上位机程序的实现 .....	14
4.4 近距离 WIFI HaLow 通信方案详细设计 .....	15
4.4.1 AH8301 WIFI 模组通信测试 .....	15
4.4.2 MQTT 协议 .....	16
4.4.3 音视频实时同传技术的实现 .....	17

4.5 远距离通信方案详细设计 .....	18
4.5.1 UDP 打洞技术 .....	18
第5章 系统测试与分析 .....	20
5.1 音视频传输测试 .....	21
5.2 传感器数据和遥控命令传输测试 .....	21
第6章 总结 .....	22

# 第 1 章 作品概述

## 1.1 作品意义

随着科学技术的不断进步，特种机器人以不同的应用方式越来越得到大家的认可。机器人设计是一个多学科交叉融合的领域，通常情况下，救援机器人可以很好地适应使用环境，甚至可以达到高度智能化程度，在代替救援人员深入危险环境进行救援方面有着很大的优越性。所以将特种机器人应用到灾害救灾领域是必需的，也是可行的。灾后救援机器人可以代替救援人员深入废墟危险环境成为灾后辅助救援的有效工具之一，在灾后使用救援机器人进行作业具有以下非常明显的优势：

- (1) 可以连续执行救援任务。
- (2) 可以在高温、高压等危险环境下工作。
- (3) 可以搭载相应的工具深入废墟空间进行远程操控救援。
- (4) 可以深入救援人员无法到达的地方进行搜索与救援。
- (5) 重量相对较轻，可以减少不稳定废墟结构二次坍塌的危险。

虽然救援机器人因具有可以连续执行救援任务，代替救援人员深入危险环境，可以搭载相应的工具深入废墟环境进行救援得到肯定与使用，但是救援机器人依然存在着受地震废墟环境影响。由于很多灾后现场的供电设施和通信基站处于瘫痪状态，导致目前通用 4G 网络在灾后地区处于微弱信号或无法使用的情况。这种通信不稳定以及动态的环境下进行，会进一步导致救援机器人在不确定性的通信条件下进行作业，从而会大大影响灾后救援作业的效率。因此提高机器人系统在通信不稳定的情况下能正常与主控台进行通信显得尤为重要。

综上所述，不论是当前救援对救援机器人的需求，还是机器人技术向前发展的推动作用，我们都需要对应急通信情况下救援机器人进行设计研究，不仅是灾后发生后辅助救援、稳定社会发展和保障人民生命财产安全的需要，还是经济发展的需要，也是救援机器人发展的必然趋势。

## 1.2 作品简介

本团队设计的应急通信救援机器人可选择两种通信模式，一种是正常的 4G 网络通信，该通信模式距离无限制。一种是在无法使用 4G 网络的时候，可以基于 AH8301 WIFI 模组进行 1 公里以上的自由通信，不被基站所限制，可以满足所有灾后现场的特殊通信情况，如图 1-1 所示。

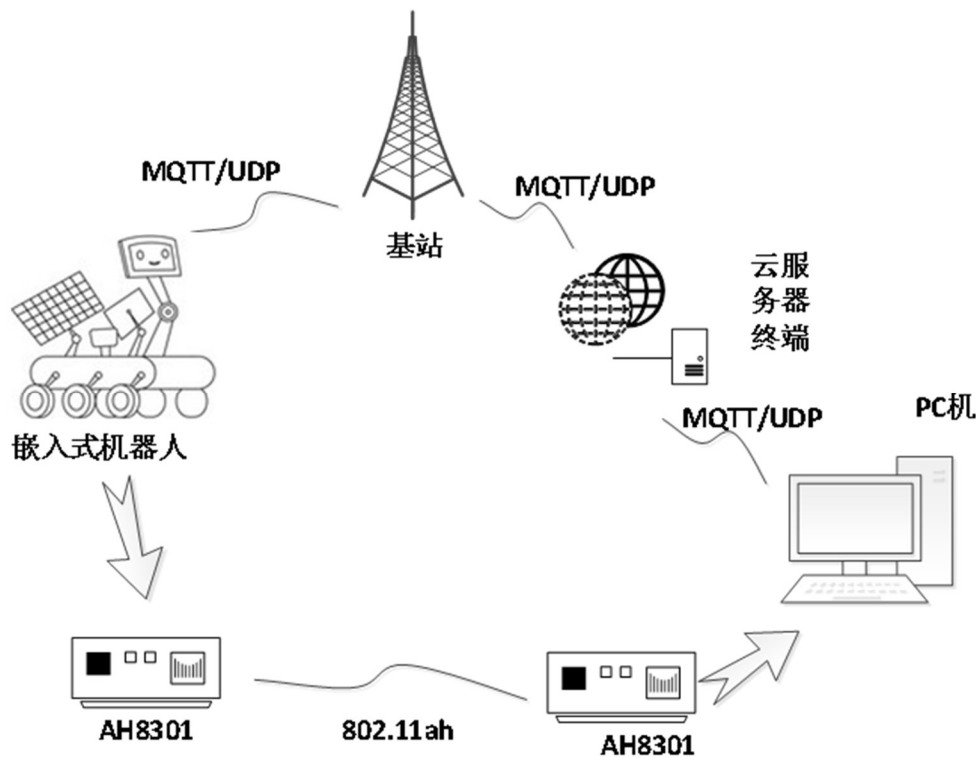


图 1-1 模式选择

除此之外，本团队设计的应急通信救援机器人可在灾后特殊环境搜救被困人群，还可以将物资精准运送到指定地区，并识别身份信息，进行物资的分发。本救援机器人还具有地图定位功能，能在第一时间将被困人员的具体位置发送给救援人员。同时还可以将采集到的音视频数据和一些传感器数据如温湿度、气体检测数据通过 AH8301 WIFI 模组或者 4G 传输给上位机，上位机也可以远程遥控机器人的运动与现场进行音视频通话。为了实现视频的实时处理，通过采用 RISC-V E902 开源处理器挂载加速 IP 实现音视频数据的实时处理，为了实现上位机和机器人之间的高效远程数据传输。

本团队在硬件设计部分完成了 G. 711 音频编码加速 IP，IP 的接口设计为通用的 AXI 总线，并通过 AHB 转 AXI IP 核将硬件加速 IP 挂载到超低功耗的玄铁 E902 处理器的 AHB 总线上。

本团队在软件部分主要完成了以下四部分工作，将 E902 作为主控平台、OK6818 开发板作为音视频数据传输媒介、STM32 作为机器人控制端媒介以及采用 QT 技术编写的上位机软件。E902 平台利用 VS1053 采集音频数据并通过 G. 711 编码 IP 编码，编码后的音频数据和视频处理的结果传送 OK6818 开发版，同时



OK6818 开发板连接了 MSK-264 以实现实时视频数据的采集，采用 RTP 流媒体传输协议将音视频数据和其它传感器数据通过 AH8301 WIFI 模组或者 4G 传输给上位机，上位机采用 QT 编写，可以实现音视频的实时播放与记录，由于 E902 的引脚数目有限，采用 STM32 作为控制端媒介，将上位机发送的控制命令通过 E902 传送给 STM32，以达到使用上位机控制救援机器人的目的，其整体系统框架图如下图 1-2 所示。

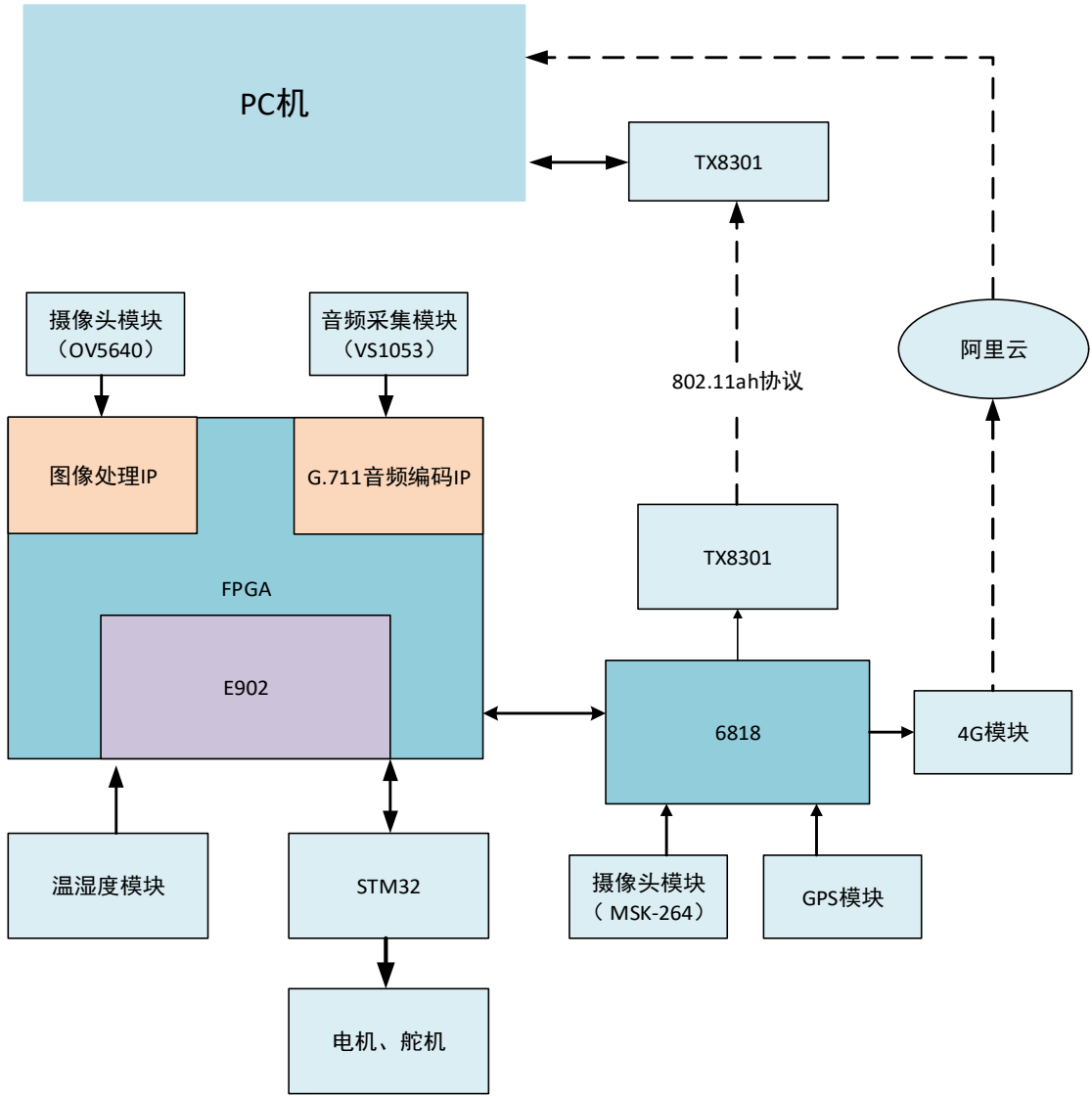


图 1-2 系统实现架构

## 第 2 章 作品难点与创新

### 2.1 作品难点

救援机器人要求在特定灾后环境中能完成救援任务,这需要救援装置能够在弱信号或者无 4G 基站情况下进行通信,如何在复杂情况下进行通信,并且定位机器人是我们团队首要解决的问题。选用了超低功耗的 RISC-V 处理器挂载加速 IP 完成设计,在 RISC-V 处理器上面完成开发难度较大。

### 2.2 关键技术研究

#### 2.2.1 实时监控信息通信技术研究

为了达到实时数据传送的要求,我们采用两种通信方式。

近距离通信传输采用内置 802.11AH 芯片的 AH8301 WIFI 模组,可以实现 1km 以上数据的实时传输,在这种传输模式下,音视频数据采用传统的 RTP 协议进行,在 OK6818 开发板搭建 MQTT 服务器,利用基于 TCP 协议的 MQTT 技术进行制命令和传感器命令订阅与发送。

远距离传输采用 4G 技术连接到云服务器,在云服务器上安装 MQTT 服务器,实现上位机和机器人之间的控制命令和传感器命令的订阅与发送,对于音视频数据,我们在云服务器部署 Node.js 环境以运行 JavaScript 程序实现 RTP 数据的转发,实现音视频数据的云服务器中转,在 4G 模式下,可以实现任意距离的数据实时传输。

#### 2.2.2 音视频实时同传技术研究

首先在 OK6818 开发板作为音视频数据的传输媒介,我们在 E902 处理器上挂载网口 IP 核,利用网口完成 OK6818 和 E902 之间的数据交互,并在 OK6818 上移植摄像头驱动,并测试摄像头采集视频的性能和视频图像质量。然后使用 H.264 编码技术对摄像头采集的视频数据进行编码压,以减少传输时所需要的网络带宽。之后基于 RTP 协议对音视频数据进行分包和组帧实现实时传输。

最后,在操控端上的图形界面程序中基于 VLC 库开发音视频播放模块,实现 RTP 音视频流媒体的接收、解码以及播放。

#### 2.2.3 音频编码的硬件加速器 IP 的设计

在超低功耗的 RISC-V 开源处理器上面通过 AHB 总线挂载音频编码硬件加速 IP 完成音频编码的硬件加速。

#### 2.2.4 火焰识别算法的硬件加速器 IP 设计

设计了 OV5640 摄像头驱动模块，使摄像头模块输出 RGB565 格式的图像数据，将 RGB565 格式的图像数据转换为 YCbCr 格式的图像数据，并通过大量的实验计算得到了，火焰的图像数据在 YCbCr 中的关系式，通过对摄像头实时采集到的图像数据进行运算，并将识别到的火焰大小发送给上位机。

## 2.3 作品创新

#### 2.3.1 RISC-V SOC 设计

与人们所熟知的 x86 和 ARM 两种架构相比，RISC-V 因其超低功耗，完全开源、可扩展性的特点受到了热捧，也是我国在处理器芯片领域实现弯道超车的最佳选择，因此我们本次设计的主控页选择在开源 RISC-V 处理器上面进行开发，并对音频编码做了硬件加速设计，改变了音频数据传输的质量，并且 RISC-V 有着开源，超低功耗的极大优点。

#### 2.3.2 无线高清音视频实时同传技术

基于 H.264 视频编码技术和音频编码技术将机器人采集的音视频数据压缩到合适的大小，利用 RTP 实时传输协议在 4G 网络下实现了音视频流媒体数据传输，在带宽有限的网络条件下，达到了较高的音频、图像质量和较低的网络时延。

#### 2.3.3 高可靠远程遥控机器人数据传输技术

在 4G 传输过程中，在云服务器上基于 JavaScript 和 Node.js 开发 UDP 打洞和转发程序，实现将遥控机器人的音视频数据转发至操控端。针对远程操控的需求，采用 MQTT 即时通讯协议和传感器数据融合技术，实现了机器人和操控端之间传感器数据和遥控命令的传输，在网络较差的情况下依然能够保证数据传输的可靠性和低延迟，改善了操控人员对机器人的操控感。

## 第 3 章 硬件设计与流程

### 3.1 RISC-V SOC 设计

本设计在开源 RISC-V IP E902 的基础上完成，将设计的硬件加速 IP 挂载在处理器的 AHB 总线上面，实现软硬件的联合开发，RISC-V SOC 的整体设计架构图如下所示：

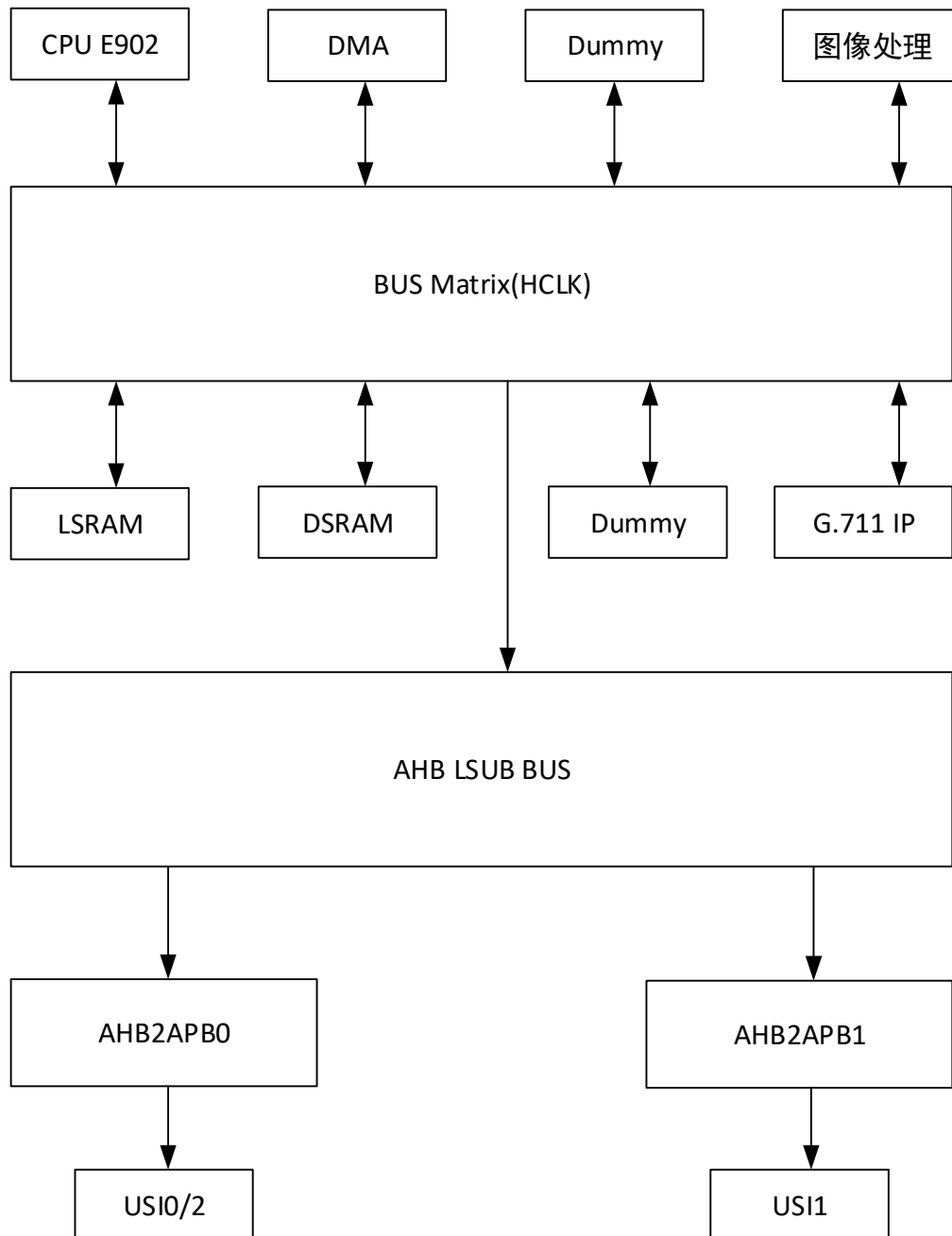


图 3-1 RISC-V SOC 整体架构图

### 3.2 G. 711 音频编码加速 IP

本设计的设计思路是将音频模块采集回来的音频数据通过 AXI 总线写入写 FIFO 中,然后将数据传输到 G. 711 编码模块进行编码,编码后将数据存入至 BRAM 中,每存 256 组音频数据后就将数据通过 AH8301 发送至上位机,其结构框图如图 4-2 所示。

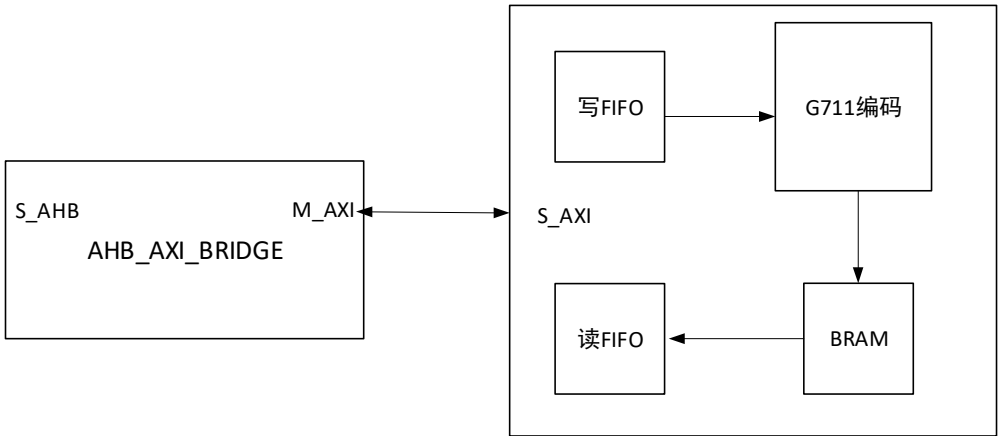


图 3-2 G. 711 音频编码 IP 框图

#### 3.1.1 G. 711 编码原理分析

G. 711 编码它是国际电信联盟 ITU-T 定制出来的一套语音压缩标准，它代表了对数 PCM (logarithmic pulse-code modulation) 抽样标准，主要用于电话。它主要用脉冲编码调制对音频采样，采样率为 8k 每秒。它利用一个 64kbps 未压缩通道传输语音讯号。起压缩率为 1:2，即把 16 位数据压缩成 8 位。G. 711 是主流的波形声音编解码器。

G. 711 标准下主要有两种压缩算法。一种是 u-Law(u-Low, u-Law, mu-law)，主要运用于北美和日本；另一种是 a-Law，主要运用于欧洲和世界其他地区。

G. 711 的内容是将 14bit (uLow) 或者 13bit (aLow) 采样的 PCM 数据编码成 8bit 的数据流，播放的时候在将此 8bit 的数据还原成 14bit 或者 13bit 进行播放，不同于 MPEG 这种对于整体或者一段数据进行考虑再进行编解码的做法，G. 711 是波形编解码算法，就是一个 sample 对应一个编码，所以压缩比固定为：

$$\begin{aligned} \frac{8}{14} &= 57\% && \text{(u-Low)} \\ \frac{8}{13} &= 62\% && \text{(a-Low)} \end{aligned} \tag{1}$$

算法原理：

a-Low 的公式如下，一般采用  $A=87.6$ 。

$$F(x) = \text{sgn}(x) \begin{cases} \frac{A|x|}{1 + \ln(A)}, |x| < \frac{1}{A} \\ \frac{1 + \ln(A|x|)}{1 + \ln(A)}, \frac{1}{A} \leq |x| \leq 1 \end{cases} \quad (2)$$

绘制函数图像，如图 3-3 所示，用  $x$  表示输入的采样值， $F(x)$  表示通过 a-Low 变换后的采样值， $y$  是对  $F(x)$  进行量化后的采样值。

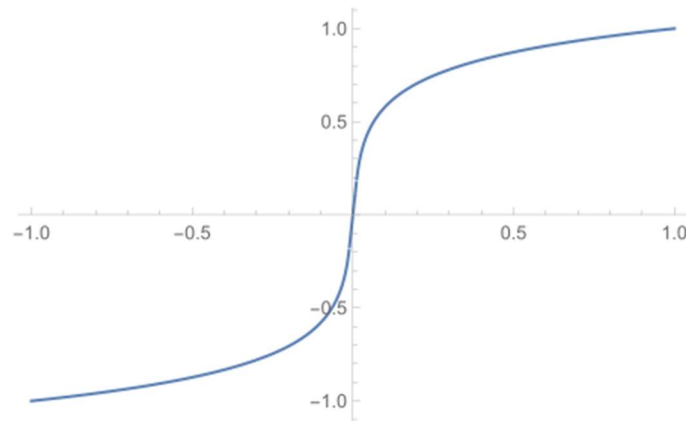


图 3-3 函数图形绘制

由此可见在输入的  $x$  为高值的时候， $F(x)$  的变化是缓慢的，有较大范围的  $x$  对应的  $F(x)$  最终被量化为同一个  $y$ ，精度较低。相反在低声强区域，也就是  $x$  为低值的时候， $F(x)$  的变化很剧烈，有较少的不同  $x$  对应的  $F(x)$  被量化为同一个  $y$ 。意思就是说在声音比较小的区域，精度较高，便于区分，而声音比较大的区域，精度不是那么高。

对应解码公式(即上面函数的反函数)：

$$F^{-1}(y) = \text{sgn}(y) \begin{cases} \frac{|y|(1 + \ln(A))}{A}, |y| < \frac{1}{1 + \ln(A)} \\ \frac{\exp(|y|(1 + \ln(A))^{-1})}{A}, \frac{1}{1 + \ln(A)} \leq |y| \leq 1 \end{cases} \quad (3)$$

## 3.2 火焰识别算法加速 IP

本设计的设计思路是通过 OV5640 摄像头驱动模块将 OV5640 摄像头采集到的数据通过写 FIFO 模块 WRFIFO 写入 DDR 中，然后通过读 FIFO 模块读出，读出的数据进入图像处理模块，在图像处理模块内部图像数据先由 rgb2ycbcr 模块将 RGB 转化为  $YC_bC_r$ ，然后通过实验得到的火焰色域关系式对其进行二值化处理，得到二值图像，并计算出火焰图像占正张图像的大小，将结果通过 AXI 接口发送出去。

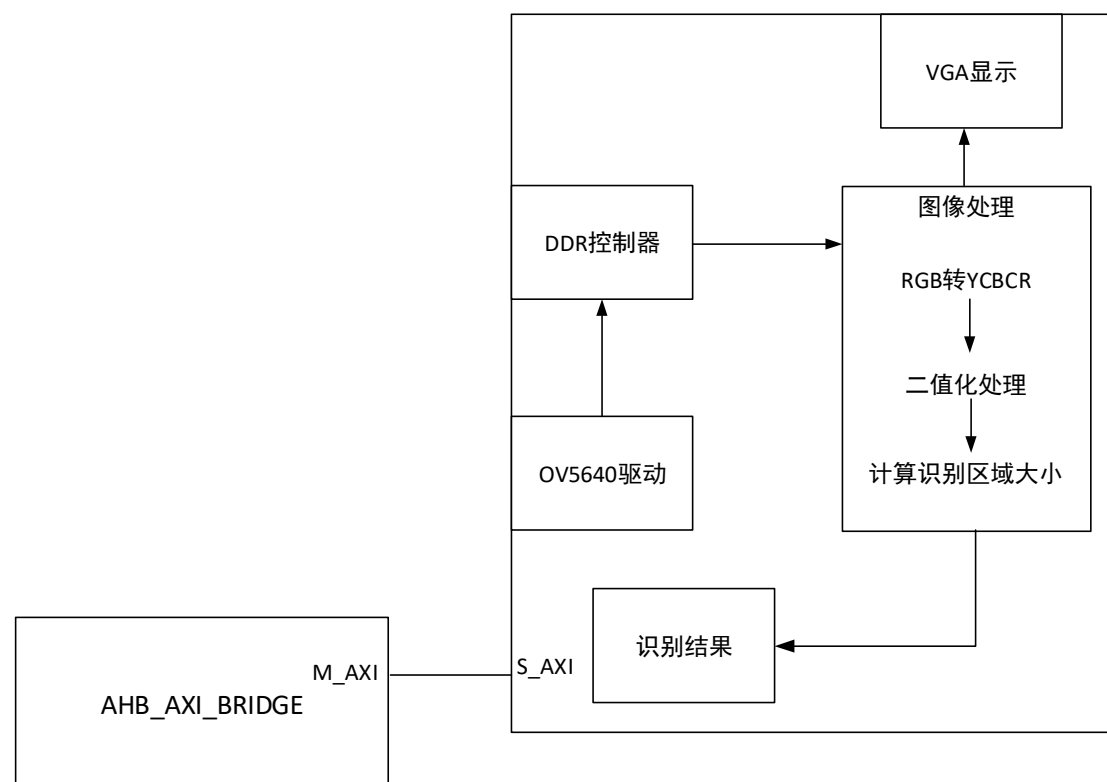


图 3-4 火焰识别 IP 框图

### 3.2.1 识别算法原理分析

$YC_bC_r$  颜色空间是广泛应用于数字视频的一种颜色空间，也是很多视频压缩编码（如 JPEG，MPEG 等标准）普遍采用的颜色表示格式。 $YC_bC_r$  颜色空间用单个分量  $Y$  来表示亮度信息，用 2 个色差分量  $C_b$ （蓝色分量）和  $C_r$ （红色分量）来表示色彩信息。 $YC_bC_r$  颜色空间和 RGB 颜色空间相互之间的转换如公式(4)所示

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.2568 & 0.5041 & 0.0979 \\ -0.1482 & -0.2910 & 0.4392 \\ 0.4392 & -0.3678 & -0.0714 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} \quad (4)$$

$YC_bC_r$  颜色空间与人类对自然界色彩的感知认识过程原理相似，且  $YC_bC_r$  颜色空间具有与 HSI 等颜色空间类似的优点：可以把色彩中的亮度信息分离出来。同时  $YC_bC_r$  颜色空间与众多硬件支持的 RGB 颜色空间的相互转换是一种线性关系，因此，分离出来的亮度信息  $Y$  并不是完全独立于色度信息。与 HSI 等颜色空间相比， $YC_bC_r$  颜色空间的空间坐标表示形式和计算过程都相对简单些。在  $YC_bC_r$  颜色空间进行火焰提取时可以充分利用火焰的色度信息和亮度信息，因此，考虑在  $YC_bC_r$  颜色空间对火焰像素进行检测。

### 3.2.2 识别算法实现

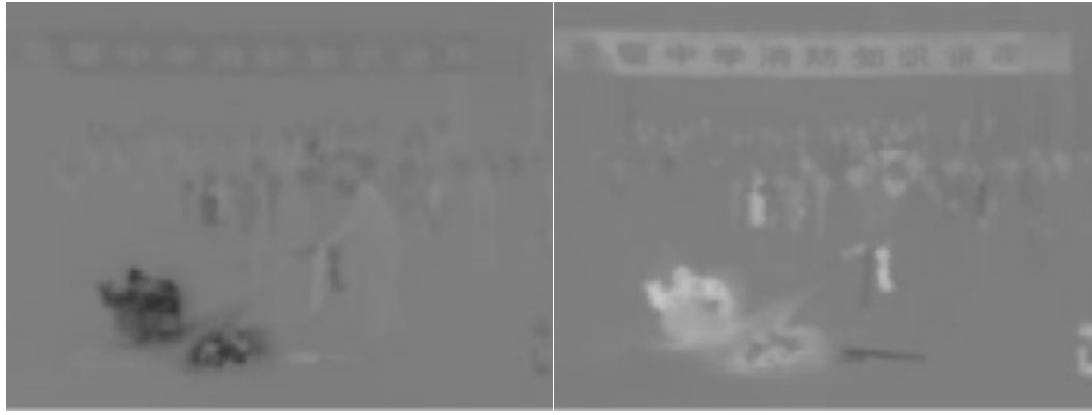
$YC_bC_r$  颜色空间和 RGB 颜色空间之间的线性转换如公式 (4) 所示。 $Y$  是亮度信息， $C_b$  和  $C_r$  是色差， $Y$  的取值在  $[16, 235]$  和的取值在  $[16, 240]$ 。为了分析火焰像素在  $YC_bC_r$  颜色空间的分布规律，本文在颜色空间对不同场景下的多张火焰图像进行  $Y$ 、 $C_b$ 、 $C_r$  三分量的分离。如图 3-5 所示。



(a) 原图像

(b)  $Y$  通道





(c)  $C_b$  通道

(d)  $C_r$  通道

图 3-5

从以图 3-5 可以看出:火焰区域的  $Y$  分量和  $C_r$  分量呈现白色或灰色,  $C_b$  分量呈现黑色。火焰图像  $Y$ ,  $C_b$ ,  $C_r$  颜色分离结果表明:火焰区域通常在整个视频图像中是亮度最大的区域;在  $C_b$  通道中, 火焰区域比整个视频图像其它区域要暗;在  $C_r$  通道中, 火焰区域比视频图像其它区域要亮。因此, 认为整张图像的三通道的均值  $Y_{mean}$ ,  $C_{bmean}$  和  $C_{rmean}$  包含着有用信息。

通过大量火焰像素的实验得出:火焰区域  $Y$  通道的值要比整个视频图像的  $Y$  通道均值  $Y_{mean}$  大, 而且火焰区域  $C_b$  通道要比整个视频图像的  $C_b$  通道均值  $C_{bmean}$  小, 同时火焰区域  $C_r$  通道要比整个视频图像的  $C_r$  通道均值  $C_{rmean}$  大, 即有

$$\begin{cases} Y(x, y) > Y_{mean} \\ Cb(x, y) < Cb_{mean} \\ Cr(x, y) > Cr_{mean} \end{cases} \quad (5)$$

其中,  $Y_{mean}$ ,  $C_{bmean}$ ,  $C_{rmean}$  分别为图像的亮度信息, 蓝色色差和红色色差的均值, 可以通过公式(6)求得

$$\begin{cases} Y_{mean} = \frac{1}{K} \sum_{i=1}^K Y(x_i, y_i) \\ Cb_{mean} = \frac{1}{K} \sum_{i=1}^K Cb(x_i, y_i) \\ Cr_{mean} = \frac{1}{K} \sum_{i=1}^K Cr(x_i, y_i) \end{cases} \quad (6)$$

式中  $K$  表示图像像素点总数。为了进一步分析火焰像素的规律, 对火焰图像进行人工地分割。

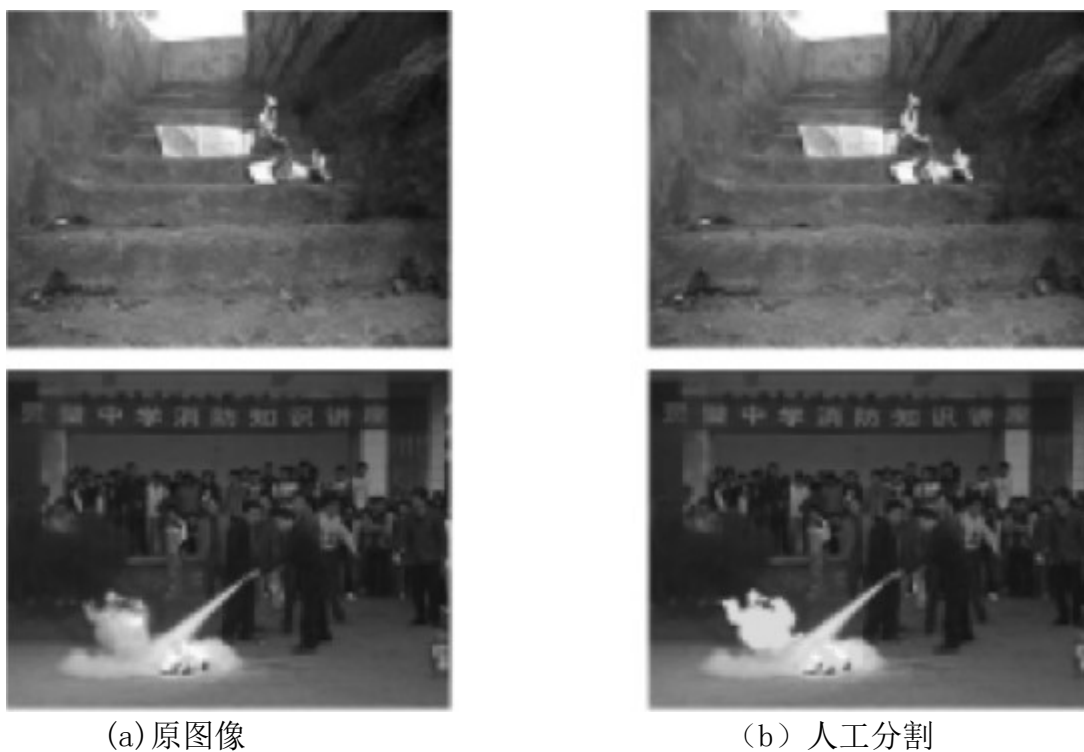


图 3-6

分别计算图 3-6 人工分割出的火焰像素的  $Y$  通道上均值  $C_r$  通道均值和  $C_b$  通道均值，计算结果如表 3-1 所示。

序号	$Y_{mean}$	$C_{bmean}$	$C_{rmean}$
1	159.83	96.098	151.09
2	163.56	82.166	165.18

表 3-1

从大量火焰图像数据可以得出在  $YC_bC_r$  颜色空间，火焰像素的  $Y$  通道的值比  $C_b$  通道的值大， $C_r$  通道的值比  $C_b$  通道的值大。火焰像素在 RGB 颜色空间满足约束条件  $R(x,y) \geq G(x,y) \geq B(x,y)$  和  $R(x,y) \geq R_{mean}$ ，转换到  $YC_bC_r$  颜色空间，则火焰像素的约束条件如式 (7) 和式 (8)。

$$Y(x,y) > C_b(x,y) \quad (7)$$

$$C_r(x,y) > C_b(x,y) \quad (8)$$

式中  $Y(x, y)$ ,  $C_b(x, y)$  和  $C_r(x, y)$  分别为像素点  $(x, y)$  的亮度分量值, 蓝色色差值, 红色色差值。

从图 1 的  $C_b$  和  $C_r$  通道可以看出: 火焰像素的  $C_b$  分量值和  $C_r$  分量值有明显不同, 火焰像素的  $C_b$  分量主要呈现“暗”色, 而  $C_r$  分量主要呈现“亮色”。于是得出火焰像素判断的另一约束条件

$$|Cr(x, y) - Cb(x, y)| \geq \tau \quad (9)$$

其中,  $\tau$  是一个设定阈值。 $\tau$  的取值为  $1 \sim 100$ , 不同的  $\tau$  值对应不同的检测率和误报率。本文通过大量实验确定  $\tau$  阈值。实验表明: 当  $\tau$  值取为 35 时, 检测率大于 96.5%, 误报率低于 15%。综上所述, 本文的火焰颜色检测算法是基于  $YC_bC_r$  颜色空间, 约束公式如式 (10) 所示。

$$\left\{ \begin{array}{l} rule1: Y(x, y) > Cb(x, y) \\ rule2: Cr(x, y) > Cb(x, y) \\ rule3: Y(x, y) > Y_{mean} \\ rule4: Cb(x, y) < Cb_{mean} \\ rule5: Cr(x, y) > Cr_{mean} \\ rule6: |Cb(x, y) - Cr(x, y)| \geq \tau \end{array} \right. \quad (10)$$

## 第 4 章 软件设计与流程

本设计的软件设计主要分为如下部分：E902 为主控平台、OK6818 开发板作为音视频数据传输媒介、STM32 作为机器人控制端媒介以及采用 QT 技术编写的上位机软件四大部分组成。E902 平台利用 VS1053 采集音频数据并通过 G.711 编码 IP 编码，利用 OV5640 采集视频数据利用图像处理 IP 处理视频数据，编码后的音频数据和视频处理的结果通过网口传送给 OK6818 开发板，同时 OK6818 开发板连接了 MSK-264 以实现实时视频数据的采集，采用 RTP 流媒体传输协议将音视频数据和其它传感器数据通过 AH8301 传输给上位机，上位机采用 QT 技术编写，可以实现音视频的实时播放与记录，由于 E902 的引脚数目有限，采用 STM32 作为控制端媒介，将上位机发送的控制命令通过 E902 传送给 STM32，以达到上位机操控救援机器人的目的，具体设计方案如下：

### 4.1 E902 CDK 工程实现

本系统采用 Wujian100 开源的 CDK 工程作为 E902 的主控端工程，官方提供的开源代码中已经封装好常用的 USART、SPI、PWM 等接口，我们可以直接通过官方的例程来使用这些接口，在音频采集中，我们采用 VS1053 语音采集模块对语音进行采集，采集频率为 16bit 8000HZ 采样，采样数据经过 G.711 编码后发送给 OK6818 开发板，对于 SPI 传输，Wujian100 平台支持的是主机模式，从机模式传输并没有实现，OK6818 的 SPI 接口需要修改底层驱动代码才能正常使用，为此，我们在 E902 上挂载了网口的 IP 核，用于 E902 和 OK6818 开发板的数据传输。OV6540 模块主要负责视频采样，我们只将结果发送给视频处理的结果发送给 OK6818。

### 4.2 OK6818 设备平台程序实现

本系统采用 OK6818 开发板作为音视频数据传输媒介，首先在 OK6818 开发板上移植 MSK-264 摄像头驱动，并测试摄像头采集视频的性能和视频图像质量。使用 H.264 编码技术对摄像头采集的视频数据进行编码压缩，以减少传输时所需要的网络带宽。之后基于 RTP 协议对音视频数据进行分包和组帧并利用 AH8301 实现实时传输。为了达到控制命令和传感器数据的高可靠性传输，我们在 OK6818 上搭建了 MQTT 服务器，通过数据订阅和发送的方式实现控制命令和传感器数据的高可靠性传输。

### 4.3 上位机程序的实现

上位机操控端为基于 Linux 操作系统的电脑，需要在上面开发友好的用户图

形界面程序来供操控人员使用，本系统选用 QT 进行用户界面程序的开发，如图 4-1 所示。QT 是一个跨平台 C++图形用户界面应用程序开发框架，其因为面向对象、跨平台、方便扩展，且允许组件编程等特性，被广泛用于用户图形界面程序的开发。最终的上位机程序可以进行音视频媒体流的实时播放和记录、各项传感器数据的显示以及利用键盘控制救援机器人运动的作用。

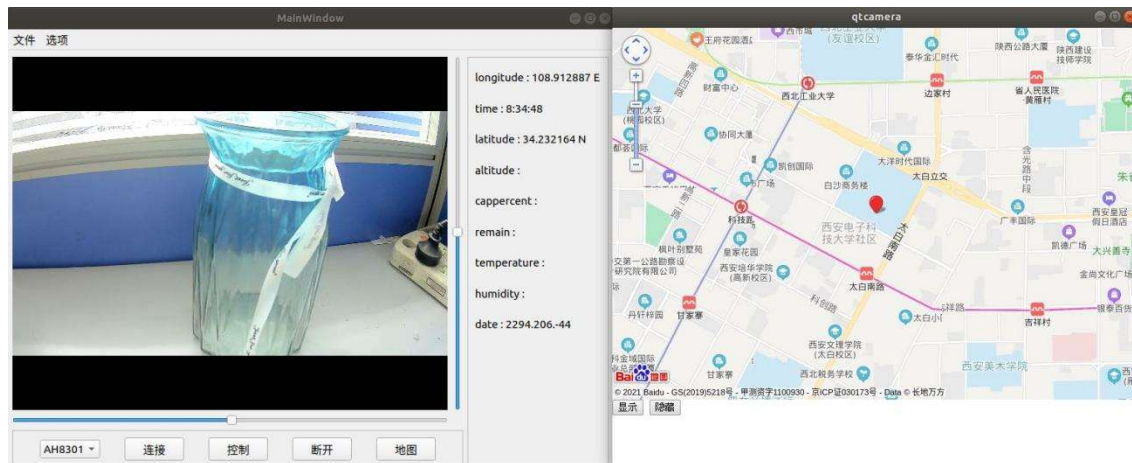


图 4-1 上位机程序

## 4.4 近距离 WIFI HaLow 通信方案详细设计

### 4.4.1 AH8301 WIFI 模组通信测试

本作品 OK6818 开发板与 PC 端的近距离通信使用了平头哥公司提供的 AH8301 WIFI 通信模组，该模组集成了某公司的一款 802.11ah 芯片，该芯片是全球第一颗 802.11ah 量产芯片，具有穿透性强、传播距离远的优点，最大传播距离可达 2KM，其中 PC 端所连接的 AH8301 WIFI 通信模组作为服务端（AP 端），OK6818 所连接的 AH8301 WIFI 通信模组作为客户端（STA 端），具体通信测试如图 4-2 和图 4-3 所示：

```
haha@haha-OptiPlex-7060:~/work/ah8301/camera
[ 5] 36.00-37.00 sec 877 KBytes 7.18 Mbits/sec 0.313 ms 0/619 (0%)
[ 5] 37.00-38.00 sec 874 KBytes 7.16 Mbits/sec 0.403 ms 0/618 (0%)
[ 5] 38.00-39.00 sec 875 KBytes 7.17 Mbits/sec 0.642 ms 0/611 (0%)
[ 5] 39.00-40.00 sec 897 KBytes 7.34 Mbits/sec 0.276 ms 0/625 (0%)
[ 5] 40.00-41.00 sec 882 KBytes 7.23 Mbits/sec 0.320 ms 0/618 (0%)
[ 5] 41.00-42.00 sec 875 KBytes 7.17 Mbits/sec 0.338 ms 0/618 (0%)
[ 5] 42.00-43.00 sec 872 KBytes 7.15 Mbits/sec 0.425 ms 0/617 (0%)
[ 5] 43.00-44.00 sec 885 KBytes 7.25 Mbits/sec 0.334 ms 0/618 (0%)
[ 5] 44.00-45.00 sec 867 KBytes 7.10 Mbits/sec 0.558 ms 0/613 (0%)
[ 5] 45.00-46.00 sec 888 KBytes 7.27 Mbits/sec 0.497 ms 0/621 (0%)
[ 5] 46.00-47.00 sec 908 KBytes 7.44 Mbits/sec 0.335 ms 0/617 (0%)
[ 5] 47.00-48.00 sec 887 KBytes 7.26 Mbits/sec 0.621 ms 0/617 (0%)
[ 5] 48.00-49.00 sec 892 KBytes 7.31 Mbits/sec 0.545 ms 0/619 (0%)
[ 5] 49.00-50.00 sec 885 KBytes 7.25 Mbits/sec 0.322 ms 0/619 (0%)
[ 5] 50.00-51.00 sec 887 KBytes 7.26 Mbits/sec 0.741 ms 0/613 (0%)
[ 5] 51.00-52.00 sec 888 KBytes 7.27 Mbits/sec 0.337 ms 0/622 (0%)
[ 5] 52.00-53.00 sec 904 KBytes 7.40 Mbits/sec 0.364 ms 0/619 (0%)
[ 5] 53.00-54.00 sec 884 KBytes 7.24 Mbits/sec 0.350 ms 0/617 (0%)
[ 5] 54.00-55.00 sec 888 KBytes 7.28 Mbits/sec 0.284 ms 0/619 (0%)
[ 5] 55.00-56.00 sec 878 KBytes 7.19 Mbits/sec 0.342 ms 0/617 (0%)
[ 5] 56.00-57.00 sec 882 KBytes 7.23 Mbits/sec 0.475 ms 0/616 (0%)
[ 5] 57.00-58.00 sec 877 KBytes 7.18 Mbits/sec 0.469 ms 0/619 (0%)
[ 5] 58.00-59.00 sec 868 KBytes 7.11 Mbits/sec 0.733 ms 0/614 (0%)
[ 5] 59.00-60.00 sec 887 KBytes 7.26 Mbits/sec 0.437 ms 0/623 (0%)
[ 5] 60.00-60.07 sec 59.4 KBytes 6.71 Mbits/sec 0.551 ms 0/42 (0%)
-----
[ ID] Interval      Transfer      Bitrate      Jitter      Lost/Total Datagrams
[SUM] 0.0-60.1 sec 1959 datagrams received out-of-order
[ 5] 0.00-60.07 sec 52.0 MBytes 7.26 Mbits/sec 0.551 ms 0/37102 (0%) receiver
Server listening on 5201 (test #2)
```

图 4-2 AH8301 UDP 传输测试

```
Server listening on 5201 (test #1)
-----
Accepted connection from 10.10.10.2, port 47464
[ 5] local 10.10.10.1 port 5201 connected to 10.10.10.2 port 47465
[ ID] Interval      Transfer      Bitrate
[ 5] 0.00-1.00 sec 782 KBytes 6.41 Mbits/sec
[ 5] 1.00-2.00 sec 857 KBytes 7.02 Mbits/sec
[ 5] 2.00-3.00 sec 844 KBytes 6.92 Mbits/sec
[ 5] 3.00-4.00 sec 854 KBytes 7.00 Mbits/sec
[ 5] 4.00-5.00 sec 848 KBytes 6.95 Mbits/sec
[ 5] 5.00-6.00 sec 856 KBytes 7.01 Mbits/sec
[ 5] 6.00-7.00 sec 856 KBytes 7.01 Mbits/sec
[ 5] 7.00-8.00 sec 847 KBytes 6.94 Mbits/sec
[ 5] 8.00-9.00 sec 847 KBytes 6.94 Mbits/sec
[ 5] 9.00-10.00 sec 856 KBytes 7.01 Mbits/sec
[ 5] 10.00-10.66 sec 560 KBytes 6.93 Mbits/sec
-----
[ ID] Interval      Transfer      Bitrate
[ 5] 0.00-10.66 sec 8.80 MBytes 6.92 Mbits/sec receiver
```

图 4-3 AH8301 吞吐量测试

从测试图中可以看出，AH8301 WIFI 模组工作在 8M 带宽模式下，传输速度可以达到 7Mb/s，UDP 传输测试中，丢包率几乎为 0，可以满足 H.264 720p 视频数据、音频数据、各项传感器数据和控制数据的传输要求。

#### 4.4.2 MQTT 协议

我们选用 MQTT 即时通讯协议作为本设计中传感器数据和控制命令传输所使用的传输协议。MQTT 协议由 IBM 公司提出，因其高效、轻量和开放的特点，它成为了物联网平台以及移动平台最受欢迎的传输协议之一。

MQTT 协议的通信模式是发布/订阅模式，该模式是通过以消息代理服务器为中介实现发布端到订阅端的通信功能，首先发布端把消息发送到消息代理服务器的消息队列中，然后消息代理服务器再从消息队列中依次转发消息到不同的订阅

端。MQTT 基本架构如图 4-4 所示。



图 4-4 MQTT 基本架构

MQTT 协议具有三种不同 QoS（Quality of Service，服务质量）的消息传输模式，如下表 4-1 所示。

QoS 值	二进制表示	消息接收次数	可靠性
0	00	至多一次	不可靠
1	01	至少一次	可靠
2	10	只有一次	可靠

表 4-1 MQTT 协议的消息传输模式

1) Level 0：最多一次的传输

消息是基于 TCP/IP 网络传输的。没有回应，在协议中也没有定义重传的语义。消息可能到达服务器 1 次，也可能根本不会到达。

2) level 1：至少一次的传输

服务器接收到消息会被确认，通过传输一个 PUBACK 信息。如果有一个可以辨认的传输失败，无论是通讯连接还是发送设备，还是过了一段时间确认信息没有收到，发送方都会将消息头的 DUP 位置 1，然后再次发送消息。消息最少一次到达服务器。SUBSCRIBE 和 UNSUBSCRIBE 都使用 level 1 的 QoS。

如果客户端没有接收到 PUBACK 信息（无论是应用定义的超时，还是检测到失败然后通讯 session 重启），客户端都会再次发送 PUBLISH 信息，并且将 DUP 位置 1。

当它从客户端接收到重复的数据，服务器重新发送消息给订阅者，并且发送另一个 PUBACK 消息。

3) level 2：只有一次的传输

在 QoS level 1 上附加的协议流保证了重复的消息不会传送到接收的应用。这是最高级别的传输，当重复的消息不被允许的情况下使用。这样增加了网络流量，但是它通常是可以接受的，因为消息内容很重要。

### 4.4.3 音视频实时同传技术的实现

我们采用了 RTP 协议实现音视频数据的传输，RTP (Real-time Transport Protocol) 是一种针对多媒体数据流的网络实时传输协议，如下图 4-6 所示。在现代生活中，TCP/IP 和 UDP 等这些网络协议在在流媒体传输领域不能达到人



们需求的时候，RTP 发挥了它的重要作用。例如在视频监控、视频会议、语音电话以及音频视频的网络传输等都应用到了 RTP 协议。OK6818 端的 RTP 协议基于 RTP 库实现。首先，创建一个 RTP 会话并对其进行初始化，其中包括配置目的地址和端口、缓冲区大小、负载类型、最大包等参数。因为一帧编码后的音视频数据的大小也往往会大于 MTU 的长度，所以我们需要对每一帧音视频数据进行分包，使单个包的大小小于 MTU 的长度，以保证 RTP 传输的性能和延迟。另外，在进行每一帧音视频数据传输时，需要根据实际时间增加 RTP 包中的时间戳，以保证在存在网络抖动的情况下，接收端也能顺利完成 RTP 包的接收与组包。

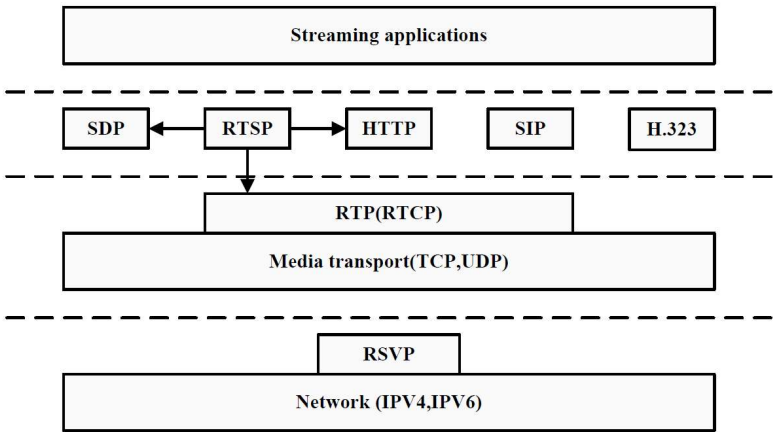


图 4-6 RTP 协议架构

## 4.5 远距离通信方案详细设计

在 4G 通信方案中，遥控机器人和操控端之间需要不断的进行数据交互，遥控机器人要向操控端上传传感器的数据，操控端向遥控机器人发送遥控命令。这就要求我们使用一个可靠的传输协议来实现以上的数据传输，同时也要保证网络时延在可接受的范围之内。因为遥控机器人和操作端都没有暴露在公网之中，所以我们选择引入云服务器，采用 C/S 的网络架构来实现即时通讯，遥控机器人和操控端都作为客服端，云服务器作为服务端，所有的数据交互经由云服务器进行转发。在云服务器上，我们采用 UDP 打洞技术对音视频数据进行转接，另外，为了保证控制命和传感器命令的高可靠性传输，我们在云服务器上也部署了 MQTT 服务器，用以控制命和传感器命令数据的高可靠性转接。

### 4.5.1 UDP 打洞技术

当今互联网到处存在着一些中间件(Middleboxes)，如 NAT 和防火墙，导致两个不在相同内网中的客户端无法直接通信。大多数中间件实现了一种非对称的通讯模型，即内网中的主机可以初始化对外的链接，外网的主机却不能初始化对内网的链接，除非经过中间件管理员特殊配置。在中间件为常见的 NAT 的情况下，内网中的客户端没有单独的公网 IP 地址，而是通过 NAT 转换，和其他相同



内网用户共享一个公网 IP。在 P2P 应用中，内网主机（客户端）需要对另外的终端（Peer）直接建立链接，但是发起者和响应者可能在不同的中间件后面，两者都没有公网 IP 地址。而外部对 NAT 公网 IP 和端口主动的链接或数据都会因内网设备未请求被丢弃掉。

在 4G 模式下，机器人和操控端 PC 往往都处于各自的局域网中，基于 RTP 协议的音视频流媒体数据没办法直接由遥控机器人直接传输至操控端，因此需要引入 UDP 打洞技术。

UDP 打洞及 RTP 音视频流媒体数据转发的网络架构如下图 4-7 所示：



图 4-7 网络架构

## 第 5 章 系统测试与分析

上位机显示界面如下图 5-1 所示：

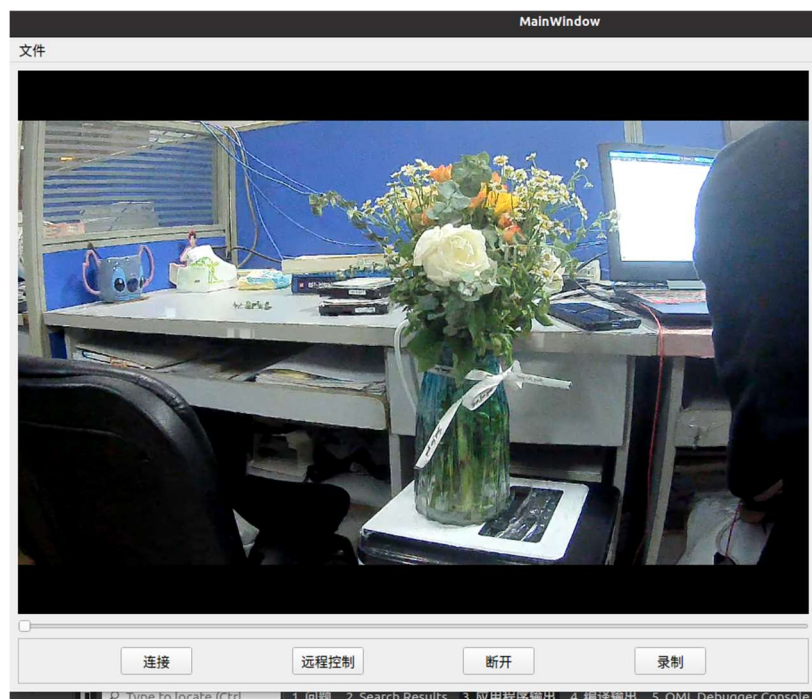


图 5-1 图像接收效果

救援机器人实物图如下图 6-2 所示：

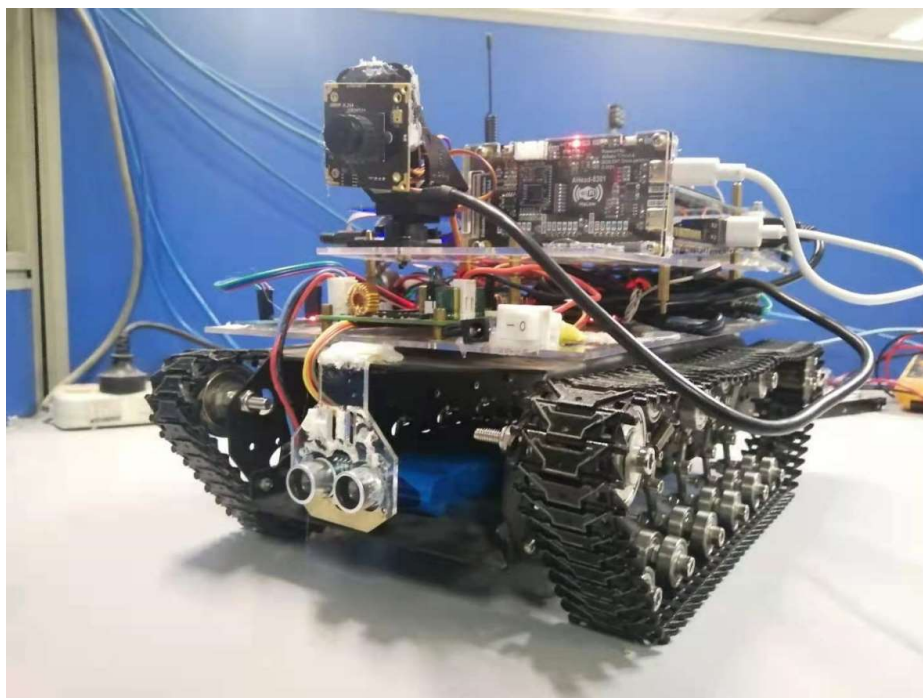


图 5-2 作品实物图

为了验证救援机器人研发中的关键技术及系统整体性能，通过搭建软硬件测试平台对项目研发结果的主要功能及性能指标进行了实际测试。

5.1 音视频传输测试

通过测试 WIFI HaLow 网络和 4G 网络下，传输各种分辨率、码率的视频数据的延时、丢包率和图像画质，从而检验实时视频传输的性能。测试结果见表。从测试结果不难看出，当视频图像分辨率越低、码率越低，视频传输的延迟越小，其原因是低分辨率、低码率的图像在编码和传输的过程中耗时更低。另外，在遥控机器人 4G 网络波动较大的情况或上行带宽较低的情况下，视频传输的延迟和丢包率都会升高。

5.2 传感器数据和遥控命令传输测试

表 6-2 MQTT 数据测试

方式	分辨率	码率	缓存值	丢包率	图像画质
4G	1080p	2M	200ms	2%	正常
		4M	200ms	3%	略有卡顿
	720p	2M	200ms	1%	正常
		4M	200ms	1%	正常
Wi-Fi HaLow	720p	2M	200ms	1%	正常
		4M	200ms	2%	略有卡顿
	480p	2M	200ms	<1%	正常
		4M	200ms	<1%	正常

表 5-1 音视频数据测试

方式	数据	延时
4G	传感器数据	60ms
	控制命令	40ms
Wi-Fi HaLow	传感器数据	40ms
	控制命令	30ms

表 5-2 传感器数据测试

## 第 6 章 总结

随着科学技术的不断进步,特种机器人以不同的应用方式越来越得到大家的认可。所以将特种机器人应用到灾害救灾领域是必需的,也是可行的。灾后救援机器人可以代替救援人员深入废墟危险环境成为灾后辅助救援的有效工具之一,在灾后使用救援机器人进行作业具有以下非常明显的优势。

虽然救援机器人因具有可以连续执行救援任务,代替救援人员深入危险环境,可以搭载相应的工具深入废墟环境进行救援得到肯定与使用,但是救援机器人依然存在着受地震废墟环境影响,由于很多灾后现场的供电设施和通信基站处于瘫痪状态,导致目前通用 4G 网络在灾后地区处于微弱信号或无法使用的情况。这种通信不稳定以及动态的环境下进行,会进一步导致救援机器人在不确定性的通信条件下进行作业,从而会大大影响灾后救援作业的效率。因此提高机器人系统在通信不稳定的情况下能正常与主控台进行通信显得尤为重要。

综上所述,不论是当前救援对救援机器人的需求,还是机器人技术向前发展的推动作用,我们都需要对应急通信情况下救援机器人进行设计研究,不仅是灾后发生后辅助救援、稳定社会发展和保障人民生命财产安全的需要,还是经济发展的需要,也是救援机器人发展的必然趋势。