

第一章 语音交互系统

1、系统功能简述

时间：20151020-20160331

HORI-Voice Interaction

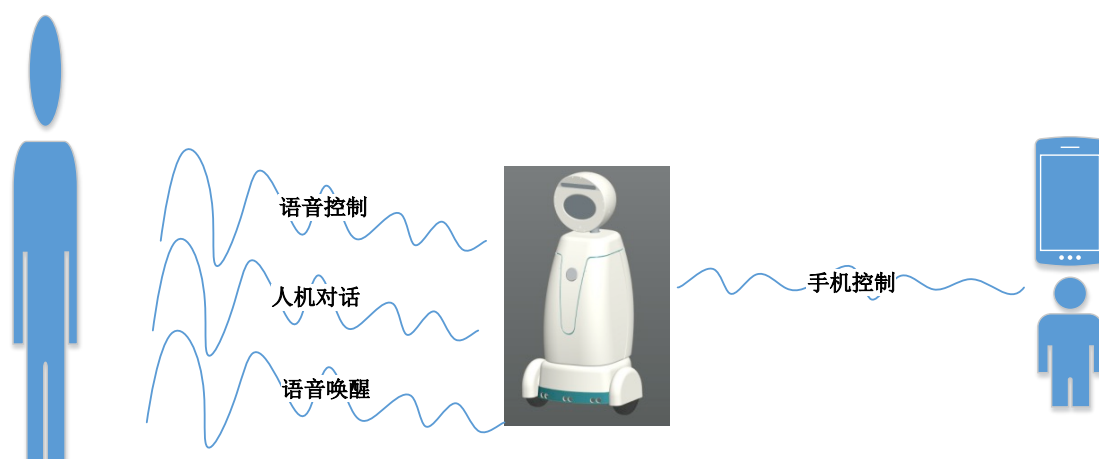


图 1.1 语音系统功能概要图

基于科大讯飞免费语音识别的 API 接口, 在机器人操作系统 ROS 下实现 HORI 机器人的以下几个功能需求:

- (1) 机器人本体的语音唤醒。
- (2) 机器人本体的语音控制。
- (3) 机器人本体的人机对话。
- (4) 接收手机端的语音数据, 实现以上三个功能。

2、系统结构简述

基于功能上的考虑以及 ROS 本身的特点, 目前语音系统主要分为以下的几个节点:

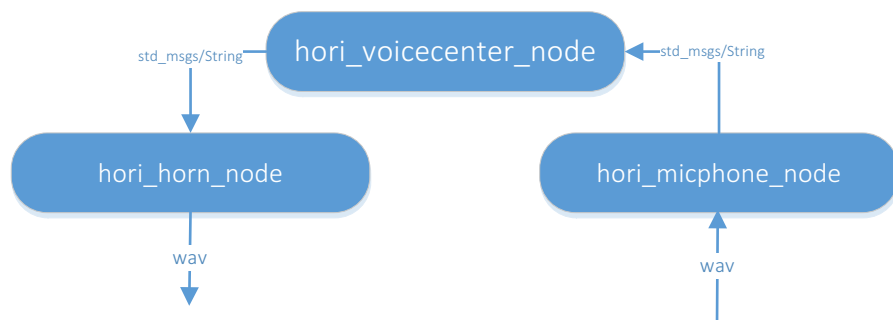


图 1.2 语音功能系统 ROS 节点

(1) hori_voicecenter_node 语音处理中心节点，功能是读取 Json 文件中的语音控制命令和基于语料库的简单问答。

(2) hori_micphone_node 语音合成节点，功能是将环境中的音频数据采集、切分，并实时保存为音频文件.wav，然后利用科大讯飞的 API 接口，实现语音到文字的转换。

(3) hori_horn_node 利用科大讯飞的 API 接口，实现文字到语音的转换。同时在 valse2016 中基于该喇叭节点可以将文本识别和纸币识别的文字结果以语音的方式表现出来。

其中，麦克风节点中涉及到了对音频信号进行采集、计算、处理的过程。其过程如下图所示：

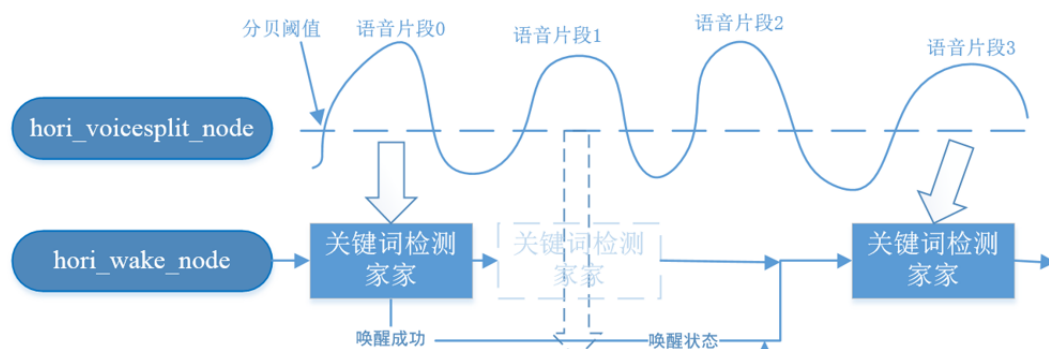


图 1.3 麦克风节点处理图

3、系统优化点

(1) 在 hori_voicecenter_node 中需要优化的是智能问答部分（应该用到自然语言处理的技术），如果采用简单的匹配方式，则需要优化语料文件的读取和语料的排序选择。

(2) 在 hori_micphone_node 节点中需要优化的部分有利用 FFT 计算分贝时，该分贝的阈值应该适应学习环境中的分贝，依据当前环境中的音频进行自适应。

(3) 在 hori_horn_node 这个部分需要优化的是代码的精简。

当然，这三个节点间的通信都是通过 ROS 话题消息的订阅和发布实现的

第二章 FPV 服务器系统

1、系统功能简述

时间：20160510-20160730

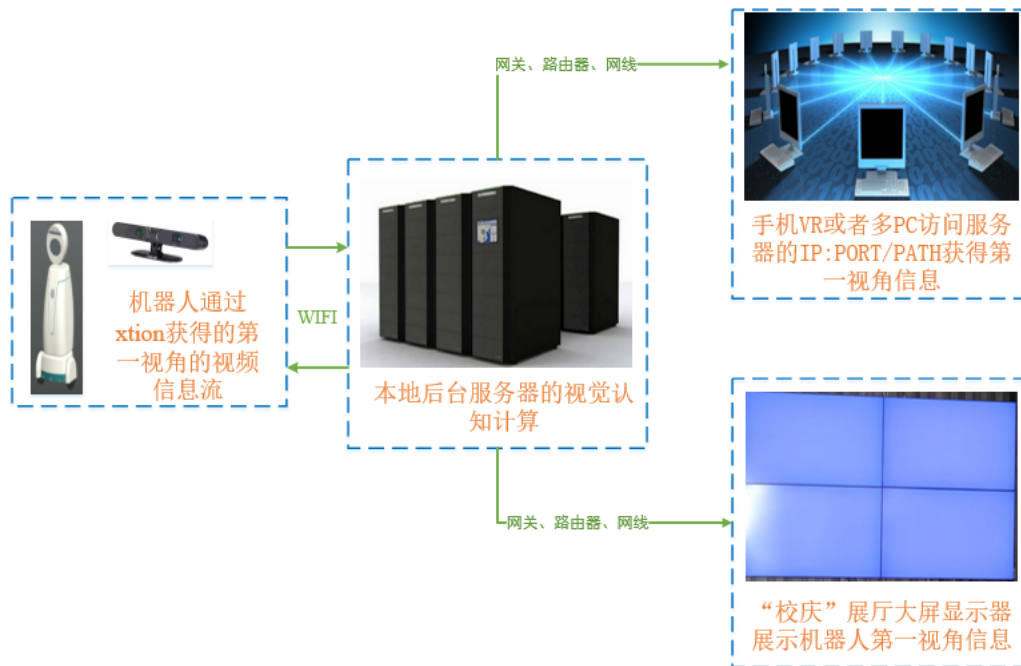


图 2.1 FPV 服务器系统

基于 Websocket 传输协议框架构建的 FPV 服务器系统的主要功能是：

- (1) 首先，在同一个局域网下，将机器人 xtion 相机获取的第一视角的 RGB 和 Depth 视频流通过 socket 套接字传输到本地后台的服务器；
- (2) 然后，基于 websocket 传输协议以及 rosbridge，将得到的 FPV 数据流打包压缩建立一个本地 web 服务器；
- (3) 最后，通过客户端访问该服务器的 IP 地址、端口和路径，便可以看到实时的 FPV 数据

2、系统结构简述

同样基于功能上的考虑以及 ROS 本身的特点，目前 FPV 服务器系统主要分为以下几个节点：

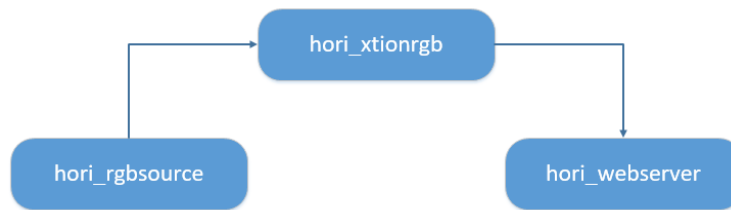


图 2.2 信号流向图

(1) hori_rgbsource 通过订阅 xtion 的话题数据，通过 socket 传输到接收端 hori_xtionrgb;

(2) hori_xtionrgb 会将得到的 FPV 数据封装并发送到一个话题/hori_rgb_image 上;

(3) hori_webserver 订阅 (2) 中的话题，然后通过 websocket 和 rosbridge 建立一个 web 服务器，这样多个客户端都可以同时访问该服务器，看到第一视角的深度图和彩色图。

这里需要说明的主要是 Websocket 异步传输协议以及 JPEG 格式数据的解码过程。在 async_web_server 中是异步传输协议的库，而 JPEG 格式数据的解码注意以 FFD8 开头以及 FFD9 结束

3、系统优化点

- (1) 深度图数据的校正，这样可以利用深度图和彩色图进行视觉算法的计算。
- (2) 深度数据和彩色数据或许可以合并成一个节点。
- (3) 基于得到的实时彩色图，其实可以在本地做一些视觉上的应用，比如物体检测，人脸识别等等。

第三章 其他小功能

1、网页首页的设计

hori_webserver 包是利用 CSS 和 HTML 设计的首页，没有涉及到与后端的交互过程，所以比较简单。首页界面如下：



图 3.1 首页界面

需要优化的点可能是界面的设计上吧，暂时就先这样了。下面的附录部分是一些系统协同开发的事情

2、Git 本地服务器的搭建

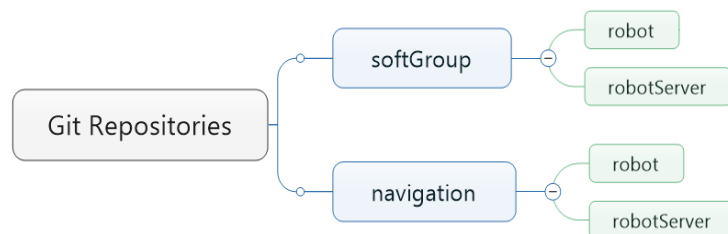


图 3.1 Git 本地仓库

附录：

时间：20160401-20160418

Valse2016 纸币识别和文字识别的移植以及与语音系统的融合工作。

时间：20160422-20160505

华为 Demo 的调试、测试和展示工作。

时间：20160801-20160929

为了校庆 116A 展厅和 2 号楼展示，协同合作开发基于前期家家机器人的校庆导览功能以及相应的测试调试工作（添加字幕功能、添加语音打断功能等）。