

Research on Voice Interaction Technology in VR Environment

Introduction

- Natural Language Processing (NLP) performance has been greatly improved by the growth of data.
- Improvements to Virtual Reality (VR) equipment have enabled higher performance and more immersive experiences.
- However, too much visual graphical noise can disturb the user's immersion. So, speech might be the best way to interact in a VR environment.
- **Goal:** To study the interactive combination of NLP, VR, and voice commands.

Development Environment

- **Hardware**

- Head Mounted Display (HMD) provides greater immersion than regular displays
 - Oculus Rift, HTC Vive, HoloLens, Playstation VR, etc.
- The Vive was selected for the hardware platform. It provides ultra-high definition image quality and also has a built-in microphone.

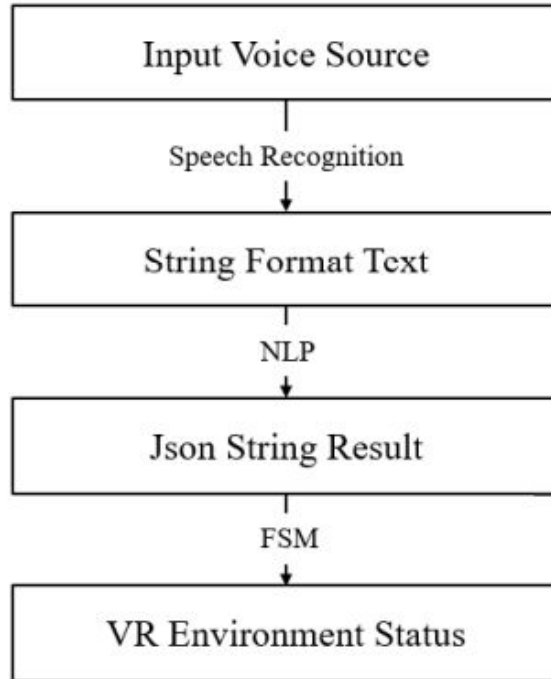
- **Software**

- Unity3D, the multi-platform game development software, was selected to create the VR environment.
- For NLP, the authors use Baidu's lexical analysis, which has functions for: word segmentation, part-of-speech tagging and named entity recognition.

System Framework

- The system is composed of three modules:
 - Voice recognition
 - NLP
 - Finite State Machine (FSM)
- The flow of the system is as follows:
 1. Monitor external audio information
 2. Convert audio signal to text
 3. Call NLP module and classify the words into a json string
 4. FSM will parse the json string and convert it into commands for the VR environment

System Framework



Function Modules

- **Speech Recognition**
 - Responsible for converting voice input into text.
 - Uses Windows.Speech API for voice recognition.
- **Natural Language Processing**
 - Responsible for analyzing and processing output of speech recognition module.
 - It creates a json string which contains information corresponding to each part of speech.
- **Finite State Machine**
 - Responsible for converting the json result into a command and triggers the corresponding state in the VR environment.

Conclusion

- The goal was to study the combination of NLP, VR, and voice commands.
- Testing the system showed that the modules executed fast and efficiently.
- In general, the system has high accuracy and can be used in virtual reality environments with low real-time requirements.

Reference:

Research on Voice Interaction Technology in VR Environment. C. Li and B. Tang. 2019 International Conference on Electronic Engineering and Informatics (EEI). 2019.