

Trabajo Final de Señales - Propuesta y Bibliografía

Lopez, Zacarías, Escalas

I. PROPUESTA

La propuesta consiste en reconocer patrones de voz y, en base a ellos, ejecutar los comandos de un videojuego. Como propuesta propia, procederemos de forma de implementar los conocimientos obtenidos en la materia y luego, aplicaremos técnicas encontradas en los distintos artículos científicos encontrados como [1], [2], [3]. Muchos de ellos nombran siempre los mismos métodos, por lo que nos centraremos en ellos más adelante.

En resumen, el reconocimiento de los comandos de voz predeterminados como ser 'arriba', 'abajo', 'atrás' y 'adelante' pueden realizarse captando la voz desde un micrófono. Una vez captada la señal, se debe reconocer el patrón, clasificarlo y ejecutar el comando asociado.

También debemos tener en cuenta que puede existir ruido, el juego puede poseer sonido propio y, además, se debe considerar el ruido ambiental al momento de jugarlo.

Una propuesta puede ser realizar un 'entrenamiento' de usuario, al inicio como menciona [2]. Pedir al usuario que diga unas pocas veces los comandos para registrarlos. Debería hacerlo bien cerca del micrófono y sin sonido del juego, para mejorar la captación. Esos patrones servirán para el posterior procesamiento en tiempo real. También podemos comparar este método con el procesamiento en tiempo real del juego sin hacer el previo guardado de algunas señales.

Guardar un patrón previo también puede ayudar a quitar los ruidos que posiblemente se encuentren en la señal como se mencionó antes. Al activar el micrófono para que el usuario grabe, podría captarse unos segundos el ruido generado por el ambiente. Este patrón servirá luego para generar un filtro que pueda quitar esas frecuencias. El ruido generado por el sonido del propio juego puede quitarse utilizando el audio con el que el juego fue desarrollado.

Cada patrón grabado por el usuario será mapeado a la correspondiente acción de usuario.

La clasificación, debe captar el micrófono y reconocer cuándo el usuario está hablando. Si es así, se realiza el filtrado de la señal captada, el reconocimiento y el mapeo correspondiente para saber qué comando ejecutar. Para tal efecto, debemos poseer un buffer donde guardar la señal. Podemos estimar su tamaño a partir de ciertas pruebas.

Se presenta un diagrama de bloques de trazo grueso para ver los módulos a gran escala de como se llevara adelante la resolución. Este esquema será la línea guía de resolución, los métodos presentados a continuación solo son formas de resolver los módulos particulares presentados en el diagrama:



Fig. 1

DIAGRAMA DE BLOQUES

II. IMPLEMENTACIÓN

Uno de los algoritmos posibles para implementar la resolución al problema es DTW [1] (Dynamic Time Warping) para obtener el parecido del comando con cada elemento del banco de comandos almacenado y MFCC (Mel Frequency Cepstral Coefficients) para extraer características de la señal de audio. Estos son los principales procesos que se le realizarían en la propuesta sobre la señal con el objetivo de obtener como resultado qué comando es el que quiere realizar el hablante. Se utilizan procesos como el ventaneo y uso de Ventanas de Hamming previos al proceso dado por el MFCC. Luego de eso usando DTW se determina si el comando recibido tiene una similitud suficiente con la base de comandos usados para comparar y se decide si tomarlo o no. Además de lo mencionado, se plantean realizar comparativas entre algunos otros métodos encontrados en la bibliografía, como es el Modelo Oculto de Markov y Redes Neuronales [3].

REFERENCES

- [1] Hui Li Xiaoling Lv Minglu Zhang. *Robot Control Based on Voice Commands*. IEEE International Conference on Automation and Logistics. Doc, 2008.
- [2] S. Sharma D. Singhal R. Gupta Y. Mittal P. Toshniwal and V. K. Mittal. *A Voice-Controlled Multi-Functional Smart Home Automation System*. Jaypee University of Information Technology, LNM Institute of Information Technology and Indian Institute of Information Technology. Doc, 1995.
- [3] P. Sukklay P. Leechor C. Pornpanomchai. *Operation of a Radio-Controlled Car by Voice Commands*. 2nd International Conference on Mechanical, Electronics Engineering - Faculty of Information, and Communication Technology, Mahidol University. Doc, 2010.