

Reconocimiento de voz humana a partir de un archivo de audio

Datos a usar y como capturarlos

Juan Andrés Vivas

24 de abril de 2018

Datos a utilizar

El conjunto de datos a utilizar serán archivos de audio con una duración de 2 segundos, formato WAV a una frecuencia de 44100 Hz.

Cantidad total requerida de datos

La cantidad total aproximada de datos requeridos será de 1000 muestras, la cual se comprenderá en un 60 % de voz y un 40 % de ruidos diferentes a la voz humana entendible.

Cantidad de datos capturados

La cantidad de datos capturados hasta el momento es de 35 muestras con ruido y 200 muestras de voz.

Consideraciones para capturar los datos

- 1 Los datos de voz se capturan con un micrófono que aislé en la medida de lo posible el ruido externo.
- 2 El ruido es capturado con un micrófono de celular.
- 3 La duración de los archivos es variable, luego sera dividida a segmentos de 2 segundos y convertida al formato requerido.
- 4 Se tomara como datos de ruido los habituales dentro de un salón o laboratorio, los cuales pueden ser murmullos, teclados de computadora, silencios, pasos, etc.

Pasos detallados para capturar los datos

- 1 Archivos de voz utilizados en los proyectos de Alvaro y Nerio.
- 2 Se captura la voz usando un audífono con micrófono y un programa de grabación hecho en python.
- 3 El ruido es capturado haciendo grabaciones con el micrófono del celular dentro del ambiente controlado, tratando de reproducir los sonidos deseados.