

UNIVERSIDAD DE COSTA RICA
Facultad de Ingeniería
Escuela de Ingeniería Eléctrica

IE0499 – Proyecto Eléctrico

Avance preliminar

I – 2020

Instrucciones

- El avance preliminar representa la formalización entre el estudiante y los profesores para trabajar en un tema determinado.
- Esta es una propuesta inicial y puede sufrir modificaciones. Aún así, el trabajo debe ser similar en alcance e intensidad.
- El documento debe ser impreso y firmado. Se entrega al coordinador del curso.

Datos generales

Estudiante

Daniel Meseguer Wong

B74767

Nombre completo

Carné

daniel@meseguer.biz

6050-2780

Correo electrónico

Teléfono

Profesor guía

Ing. Marvin Coto Jiménez, Ph.D.

marvin.coto@ucr.ac.cr

Nombre completo

Correo electrónico

Profesores lectores

Ing. Profesor Lector A

lector@ucr.ac.cr

Nombre completo

Correo electrónico

Ing. Profesor Lector B

lector@empresa.com

Nombre completo

Correo electrónico

Datos del proyecto

Título

Desarrollo de un sistema de segmentación de audios de larga duración

Descripción

La generación de corpus de habla (grabaciones con sus correspondientes transcripciones) son el insumo principal para la generación de sistemas de reconocimiento y síntesis de habla. Existen varias fuentes posibles de datos, entre las cuales se encuentran los audiolibros y las grabaciones de radio (podcast). Un primer paso necesario para la correcta utilización de estos audios, que usualmente son de larga duración, es segmentarlos, es decir, separarlos en archivos más pequeños que sean aptos para su procesamiento. La separación se realiza normalmente a partir de la detección de segmentos de silencio. En el proyecto se pretende desarrollar un sistema que pueda fragmentar de esta manera audios de larga duración, y caracterizar estadísticamente los resultados. Dado que la detección del silencio se realiza a partir de ciertos parámetros, la manipulación de éstos puede resultar en la separación efectiva o no de los diferentes segmentos. Por esto, se requiere que el sistema permita un ajuste dinámico de los parámetros, de manera que el usuario pueda seleccionarlos y evaluar los resultados a partir de la información estadística que se obtenga con el mismo sistema.

Objetivo general

Implementar un sistema basado en el lenguaje de programación Python, que permita segmentar y analizar archivos de larga duración, basándose en la detección de silencios.

Objetivos específicos

1. Valorar la importancia de las tecnologías del habla y audio, así como sus retos actuales.
2. Adaptar una implementación de segmentación automática de audio basada en la detección de silencios.
3. Establecer un sistema de análisis estadístico de los resultados (cantidad, duraciones), que permita automatizar el proceso de segmentación acorde con los fines buscados.
4. Probar la implementación del algoritmo en un conjunto de datos de prueba.

Clasificación temática

Reconocimiento de patrones

Reconocimiento y síntesis de voz

Procesamiento digital de señales

Declaración de último semestre

- ☐ Sí, este es mi último semestre.
- ☒ No, este no es mi último semestre.

Justificación: El ciclo actual no consiste en el último semestre, pues restan en el expediente académico tres cursos más a matricular el siguiente semestre del año 2020, junto con el desarrollo de la práctica profesional y el trabajo comunal universitario.

Firmas

Ing. Marvin Coto Jiménez, Ph.D.
Profesor guía

Daniel Meseguer Wong
Estudiante

Ing. Profesor Lector A
Profesor lector

Ing. Profesor Lector B
Profesor lector