

Reconocimiento de voz

Speech recognition

Autor: John Edward Ospina Ladino

IS&C, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: j.ospinal2@utp.edu.co

Resumen— Este documento presenta un análisis y posible propuesta para la solución del problema de reconocimiento de archivos de audio, correspondiente a las voces de las personas conocidas, no se presenta ni código ni pseudocódigo, es sólo una idea que con trabajo arduo se podría conseguir exitosamente

Palabras clave— voz, reconocimiento, audio, programa, redes neuronales, decibels, frecuencia

Abstract— This document presents an analysis and possible proposal for the solution of the problem of recognition of audio files, corresponding to the voices of known people, neither code nor pseudo-code is presented, it is just an idea that with hard work could be achieved successfully

Keywords— speech, recognition, audio, program, neural networks, decibels, frequency

I. INTRODUCCIÓN

La voz humana consiste en un sonido rígido emitido por un ser humano. Los pulmones deben producir un flujo de aire adecuado para que las 2 cuerdas vocales choquen entre sí.

Los sonidos producidos por la voz humana ofrece una frecuencia (hz) entre 250 hz y 3000, por este motivo se podría encontrar una solución al problema de identificación de voces de personas famosas, teniendo en cuenta que se reducirían significativamente el número de elementos a analizar debido a que no todos son figuras públicas.



Figura 1: Voz del ser humano

I.1 Planteamiento del problema

El reconocimiento de voces de personas famosas para el autor de este documento solo tiene como motivo de surgimiento la resolución de crímenes donde se vean involucradas dichas personas.

También podría ser utilizado por aficionados a la música para identificar en qué momento exactamente está cantando algún artista en específico, pero principalmente sería para identificar quién estaba hablando en un momento exacto de una llamada, así se podría generar cargos en contra o a favor de esa persona dependiendo de la situación presente.

I.2 Posible solución

El autor de este documento plantea de una manera pertinente la creación de un sistema experto que tenga en su memoria una cantidad grande de números de frecuencia enlazado con una lista de personas famosas, la primera gran diferencia que se podría plantear en el programa podría ser el sexo, de esta forma podríamos reducir considerablemente la búsqueda que realizaría el software en cada ocasión.

El sistema experto debe ser capaz de reconocer la definición de frecuencia y la importancia de esta dentro del ámbito de la resolución del problema.

Se recomienda realizar la creación de este software acompañado de personas como ingenieros de sonido o músicos que pueden dar una teoría significativa e importante al avance del mismo, debido a que ellos tienen que conocer de mejor manera los conceptos de sonido, color de voz, timbres, entre otros.

El elemento que diferencia significativamente una voz de otra cuando ya se han separado los sexos, es claramente el color de la voz, que puede ser de eufonía clara o eufonía oscura, después de minimizar aún más los posibles colores de voces, para el sistema experto sería más fácil obtener el nombre de la persona que posee esta voz.

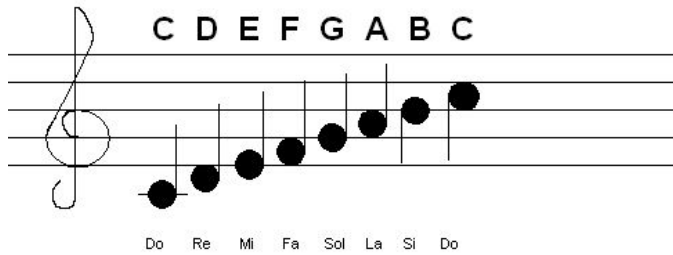


Figura 2: Notas musicales en pentagrama

También el sistema experto se puede apoyar en las notas musicales para la resolución de dicho problema debido a que las personas sean afinadas o desafinadas muy tienen cierta nota por defecto, cada una de estas notas se encuentra en una frecuencia específica.

I.2 Conclusiones

- Es posible solucionar este problema.
- Para solucionarlo se debe tener conocimiento previo en áreas como el sonido y generalidades musicales.
- Puede ser solucionado con redes neuronales, sistemas expertos, básicamente con una inteligencia artificial.