

Hito_6_Considering_the_Threshold_of_Hearing

January 8, 2025

1 Considering the Threshold of Hearing

1.1 Introducción

En este hito se tiene el objetivo de crear una implementación que utilice el Umbral de Audición Humana (ToH, por sus siglas en inglés), para mejorar la codificación y percepción del sonido. Este umbral define la sensibilidad del sistema auditivo humano en función de la frecuencia, permitiendo ajustar la cuantización del sonido en base a dicho valor.

1.2 Solución al problema

La solución teórica de este apartado se divide en dos procesos independientes pero relacionados: 1. Crear un proceso para registrar una curva ToH personalizada. Es necesario implementar un asistente mediante la consola de python que divida la banda de frecuencia audible por los humanos (de 20hz a 20000hz) en sub bandas del mismo tamaño, por ejemplo, si usamos 5 niveles, obtendremos 32 sub bandas de 625hz. Una vez divididas las sub bandas se reproduce un pulso representativo de la misma, por ejemplo la frecuencia media o las frecuencias correspondientes a los 3 cuartiles, dicho pulso comenzará con un volumen bajo y subirá progresivamente hasta que el usuario lo escuche, quedando ese volumen almacenado. El volumen puede ser procesado en porcentaje o en decibelios, por simplicidad. [nota: Nosotros en este proceso usamos el porcentaje que más tarde se convierte en db].

2. Usar la curva ToH para saber qué sub bandas deben comprimirse más : Al comienzo o durante la comunicación ha de intercambiarse las curvas ToH entre los dos usuarios para que el contrario la utilice a la hora de cuantizar. Al saber qué frecuencias el otro usuario no es capaz de percibir con facilidad, también se sabe qué sub bandas se pueden cuantizar más, pues el error será menos perceptible, es decir, se ahorrará ancho de banda sin bajar la calidad de una manera perceptible. Eso permite que las bandas que si son relevantes, tengan menos cuantización, devolviendo los kbps a su estado original y aumentando la calidad percibida.

En ambos pasos se usan wavelet packet en lugar de dwt con la intención de generar subbandas iguales en lugar de únicamente dividir una de las sub bandas en cada nivel, lo cual es necesario para que todas las sub bandas sean del mismo nivel.

Una vez los usuarios han intercambiado sus curvas ToH la comunicación sigue un proceso similar al resto de scripts vistos en el curso con la novedad de la cuantización personalizada por sub banda.

1.3 Puntos claves del código

En el diseño del sistema, se han añadido varios argumentos para facilitar la interacción y personalización de la curva ToH.

A continuación, se detalla el propósito y la funcionalidad de cada uno de ellos.

1. **-ctoh o -createToH**: Al utilizar este argumento, el usuario puede definir su propia curva ToH personalizada.
2. **-ctohc o -createToHCumulative**: Este argumento ofrece la opción de construir una versión acumulativa del ToH.
3. **-tohf o -ToHfilename**: Al utilizar este argumento se toma como valor un archivo cuyo nombre debe especificarse textualmente. Si el archivo existe, se carga la curva ToH contenida en él. Si no existe, la curva ToH creada se guarda en el archivo especificado, facilitando la reutilización.

1.4 Resultados de ejecución

Al ejecutar el código, efectivamente se le va pidiendo al usuario que interactúe dándole al intro cada vez que perciba claramente el sonido, iterando y recorriendo el intervalo de frecuencias perceptibles para el ser humano con diferentes valores en cuanto a volumen del sonido.

Dependiendo de las respuestas del usuario, se almacenarán dichos valores dentro de una estructura de datos, la cual va a representar los datos de la curva ToH personalizada para quien haga la ejecución.

1.5 Bibliografía

Para el desarrollo de esta actividad se ha hecho uso de los apuntes de la asignatura y de la documentación de las herramientas utilizadas, como Wavelet Packets.