
Capítulo 4: Escalado, filtros, bancos de filtros y espacios de representación

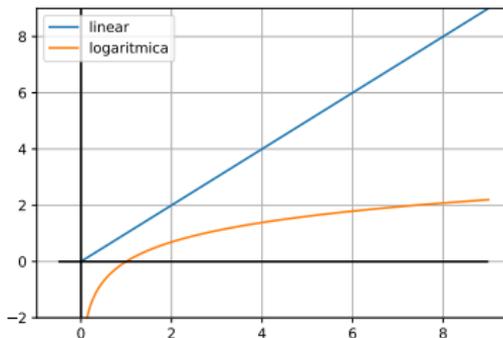
En este capítulo se presentan diversas estrategias para el análisis espectral y la construcción de representaciones que capturen aspectos relevantes de una señal. Aunque los ejemplos se centran en el dominio acústico, los conceptos desarrollados son aplicables a una amplia variedad de señales.

Escalamiento

Escalamiento en amplitud: Se refiere a como se representan las intensidades de las frecuencias en el gráfico.

Si se representa la intensidad en escala lineal, los valores grandes (como 1) dominarán la visualización. Los componentes más débiles quedarán ocultos en el fondo negro del espectrograma porque la diferencia es de varios órdenes de magnitud.

$$A_{dB} = 10\log(A_{li})$$



Colab Escalamiento

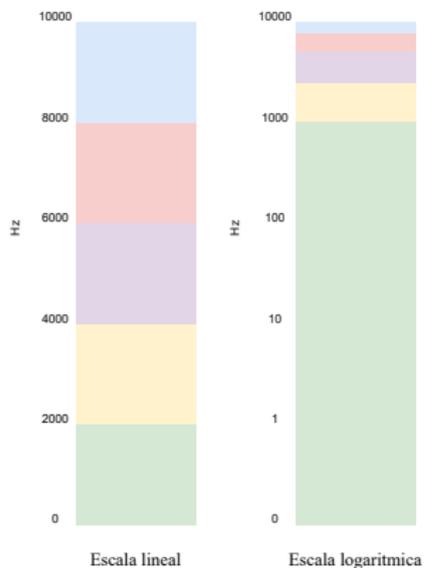
Escalamiento

Se refiere a como se representan las frecuencias en el espectrograma. Hay diferentes enfoques para escalar las frecuencias, y cada uno tiene sus propias ventajas y aplicaciones.

Escalamiento en frecuencia (escalas no psicoacústicas):

- Escala lineal: En esta escala, las frecuencias se representan en intervalos constantes. Es útil para analizar detalladamente frecuencias en un rango estrecho
- Escala logarítmica: Las frecuencias aumentan exponencialmente (no aumentan en intervalos constantes sino en múltiplos constantes.)

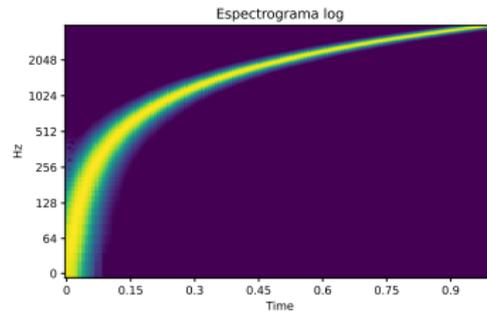
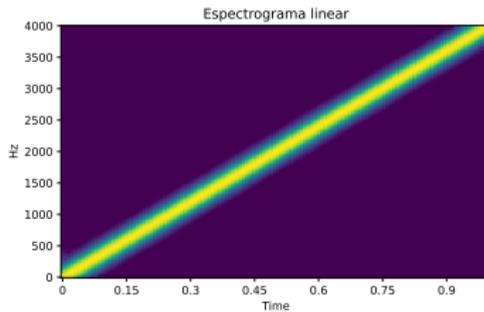
Escalamiento



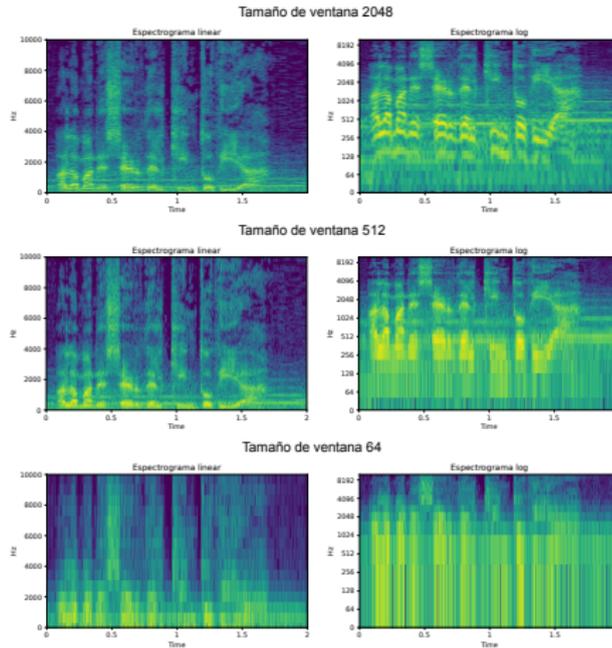
Permite enfocarse en la información que se encuentra en las frecuencias bajas (análisis de voz)

Lineal vs logarítmica

$$\cos(2\pi ft), f = 2000t\text{Hz}$$



Escalamiento y tamaño de la ventana



Escalamiento

Escalamiento en frecuencia (escalas psicoacústicas):

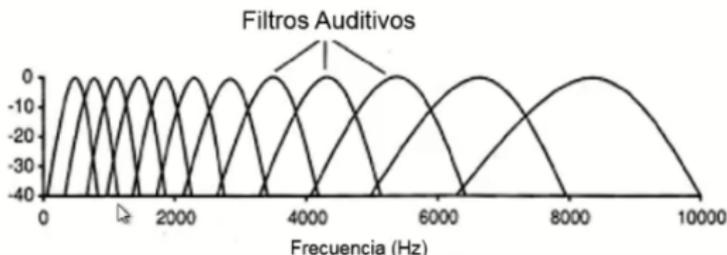
Las escalas psicoacústicas son modelos que buscan describir cómo los seres humanos perciben el sonido. Por ejemplo, la sensibilidad del oído humano a diferentes frecuencias no es uniforme; somos más sensibles a ciertas frecuencias, lo cual está relacionado con cómo se excita la membrana basilar.

[Excitación de la membrana basilar](#)

Bandas críticas del oído

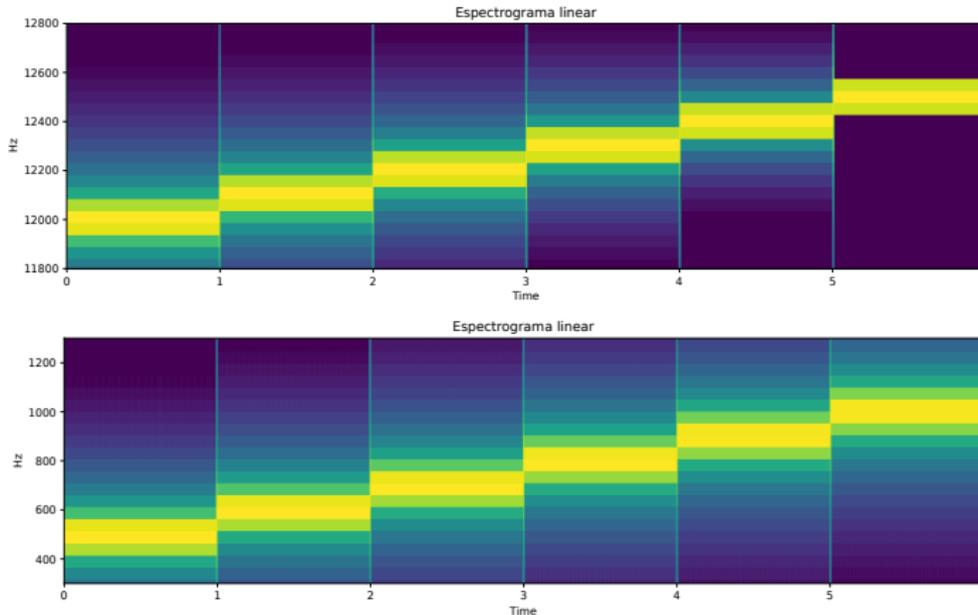
Una banda crítica es un rango de frecuencias dentro del cual los sonidos se perciben como un solo evento auditivo. Fuera de este rango, los sonidos pueden ser percibidos de manera independiente.

El ancho de una banda crítica varía con la frecuencia. Generalmente, es más estrecho en frecuencias más bajas y más amplio en frecuencias más altas. Esto significa que en frecuencias bajas, el oído puede distinguir entre sonidos más cercanos, mientras que en frecuencias altas, sonidos más distantes en frecuencia pueden ser percibidos juntos.



Es similar a un comportamiento logarítmico.

Bandas críticas del oído



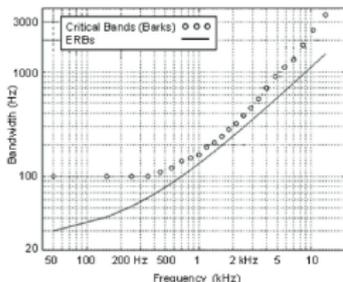
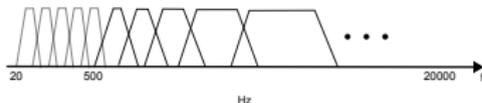
Audios de bandas críticas

Escalamiento: escalas auditivas o psicoacústicas

Esta escala busca representar incrementos de frecuencia desde un punto de vista perceptivo. Se desea que cada incremento en frecuencia sea un incremento en una banda crítica. Mediante test subjetivos es posible estimar las bandas críticas del oído humano. Lo que da lugar a diferentes escalas: Bark, ERB, Mel.

Escala de Bark

Este modelo propone que las bandas críticas tienen un ancho lineal de 100 Hz hasta 500 Hz y se incrementan de forma logarítmica desde 500 Hz

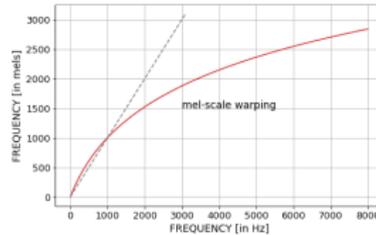


Escala de Mel

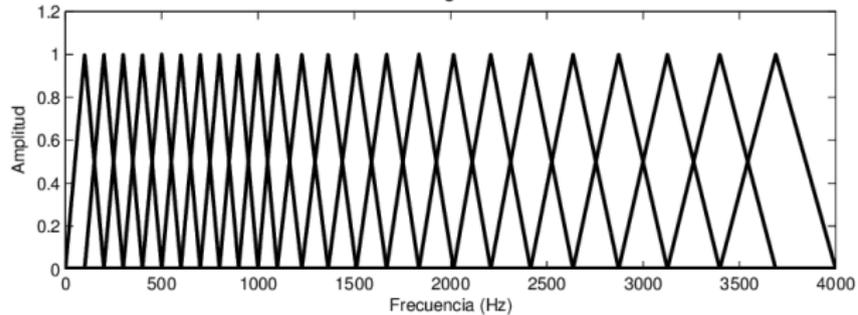
Es una de las primeras escalas psicoacústicas propuesta en 1937 por Stevens, & Newmann. Se construyó a partir de test psicoacústicos con tonos que se percibían a una misma distancia de otros (similar al ejemplo mostrado previamente). A pesar de su antigüedad es muy usada actualmente en aplicaciones de reconocimiento de voz, análisis habla patológica, entre otras. La escala completa se puede definir como:

$$2595 \log_{10} \left(1 + \frac{f_{Hz}}{700} \right)$$

Escala de Mel

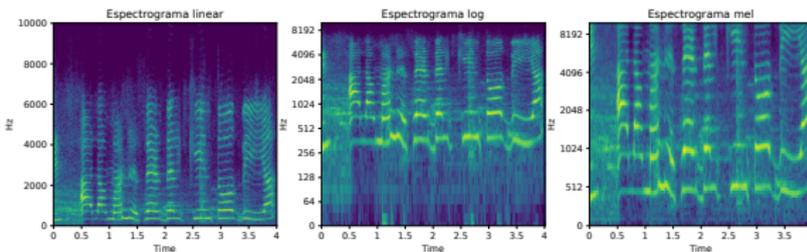
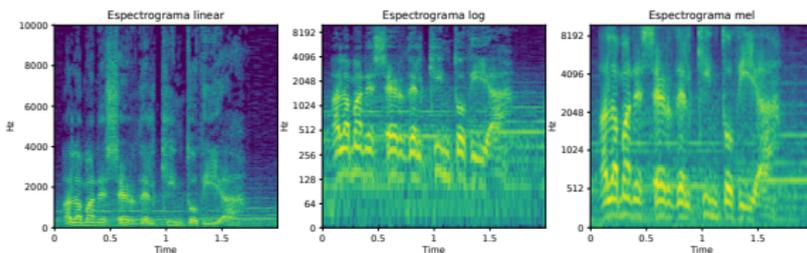


Banco de filtros triangular en la escala mel



Comparación de escalas

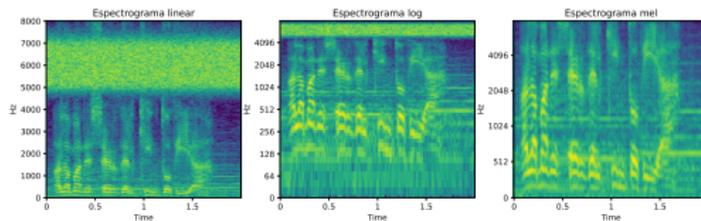
Comparación misma frase dos personas distintas



Comparación de escalas

Comparación ruido en distintas bandas de frecuencia

Ruido en alta frecuencia



Ruido en baja frecuencia

