

Tarea 7-8

Universidad Autónoma De Nuevo León
Facultad de Ciencias Físico-Matemáticas
Procesamiento y clasificación de datos

Procesamiento de audio

Nombre: Ricardo Luna Escobedo

Matricula: 1805328

Fecha: 16/03/2025

1. Resumen

En este trabajo observaremos el empleo de la programación en python para poder llevar a cabo un análisis entorno al sonido, pero desde un punto de vista que involucre a las matemáticas para el procesamiento a nivel de decibeles, frecuencias y ondas.

Siendo más objetivos, vamos a tomar 2 pistas audio con contenido musical y observaremos su composición. Las pistas elegidas son: "OH YEAH!" de **Big Time Rush**, y "Paperback Writer" de **Los Beatles**. Haremos una comparativa sobre como fue procesado/manipulado el audio de ambas canciones, ya que ambas fueron grabadas en un época donde la ingeniería audio, los géneros, los instrumentos y la manera de componer son muy distintas.

2. Introducción

Las nuevas tecnologías nos han brindado un mejoramiento en distintos ámbitos, entre ellos: **la música**. Un ejemplo de ellos es la implementación de los efectos, máquinas análogas/digitales, dispositivos de conexión inalámbrica, ¿y por qué no?, la visualización matemática de melodías. Pero... ¿nunca se han preguntado como es esto posible?, ¿cómo las matemáticas nos pueden ayudar a comprender algo que la mayoría cree solo poder estudiar/analizar auditivamente? Sabemos que si conectamos una guitarra a un amplificador, podremos escuchar su sonido potencializado, ¿pero qué sucede detrás del sonido que sale del amplificador? ¿cómo podemos conocer sus características internas? la respuesta a estas preguntas se reduce a lo siguiente: **procesamiento de audio**.

En el mundo musical, se perciben las cosas de una manera más visual y escrita. Se tienen elementos como las notas, silencios, compases, tonalidades y rítmicas. Aunque los elementos anteriores se pueden medir numéricamente, hay una rama en particular que permite visualizarlos de una manera más detallada y matemática: **La ingeniería en audio**. Es en esta parte en donde el procesamiento de audio está implícito.

3. Investigación

3.1. Datos

1. ¿Qué es el sonido?

Es la propagación de las ondas mecánicas originadas por la vibración de un cuerpo a través de un fluido o un medio elástico. Dichas ondas pueden o no ser percibidas por los seres vivos, dependiendo de su frecuencia. Éstas se desplazan expansivamente, a una velocidad promedio (en aire) de 331,5 m/s, y pueden reverberar ("rebotar") en distintos tipos de superficies, logrando distintos efectos de eco o de distorsión, que a menudo magnifican su potencia (como en las cajas de resonancia o los parlantes).

El sonido presenta las siguientes características físicas:

- **Frecuencia.** Número de vibraciones completas por segundo que efectúa la fuente del sonido y que se transmite en las ondas.
- **Amplitud.** Intensidad (potencia acústica), que solemos llamar "volumen". La amplitud se relaciona con la cantidad de energía transmitida por las ondas sonoras.
- **Longitud de onda.** Distancia que recorre una onda en un período de oscilación, o dicho de otro modo, la distancia entre dos máximos consecutivos de la oscilación.
- **Potencia acústica.** Cantidad de energía emitida por las ondas por unidad de tiempo. Se mide en vatios y depende directamente de la amplitud de onda.
- **Espectro de frecuencia.** Distribución de amplitudes, o energía acústica, para cada frecuencia de las diversas ondas que componen el sonido. [1]

2. ¿Qué es la música?

La música es una forma de expresión artística que consiste en la combinación de sonidos vocales o instrumentales, conforme a estándares culturales de ritmo y melodía. Es una de las más antiguas formas de arte y ha estado presente, de diferentes maneras, en prácticamente todas las culturas humanas. Se puede crear con fines estéticos, recreativos, ceremoniales, terapéuticos o de comunicación.

Existen muy diversas maneras de producir música. Se pueden emplear métodos tradicionales con instrumentos y voces, improvisar con objetos de uso cotidiano o utilizar programas informáticos de creación musical, llamados DAW (estación de audio digital, por sus siglas en inglés). Sea cual sea el método, el principio es el mismo.

La música depende de cuatro parámetros fundamentales del sonido, que son:

- **Altura.** Es la frecuencia de vibración del sonido. Dependiendo de cuántas veces por segundo vibre, el sonido será más grave (bajo), o agudo (alto).
- **Duración.** Es el tiempo en que permanece audible el sonido.
- **Intensidad.** La intensidad es la fuerza con la que se produce un sonido, o la cantidad de energía acústica que contiene. Se mide en dB (decibeles) y tiene relación con el volumen o la presión sonora que perciben nuestros oídos.
- **Timbre.** Es la cualidad del sonido determinada por la forma de las ondas que lo componen. Es lo que permite distinguir un sonido de otro, aunque tenga la misma altura, duración e intensidad. [2]

3.2. Descripción

1. Pistas utilizadas

1.1. Pop vs Rock

Figura 1: Grupos elegidos



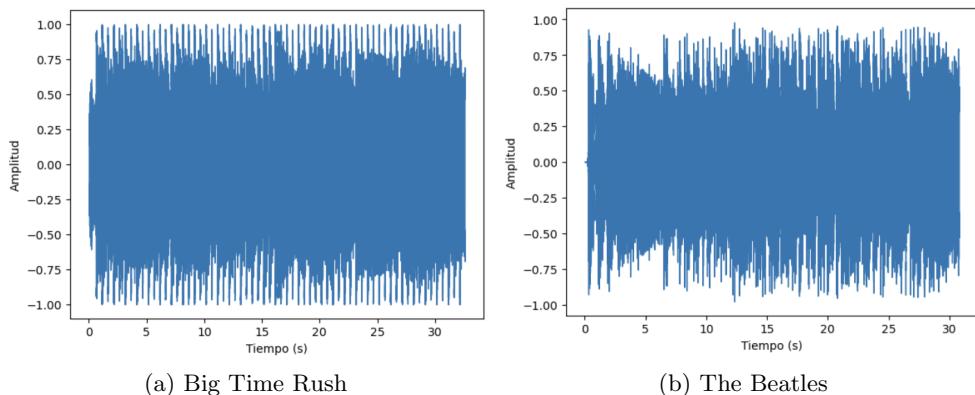
(a) Big Time Rush

(b) The Beatles

2. Análisis de amplitud

Comenzamos graficando la amplitud de las ondas de ambas pistas.

Figura 2: Amplitud de ondas



(a) Big Time Rush

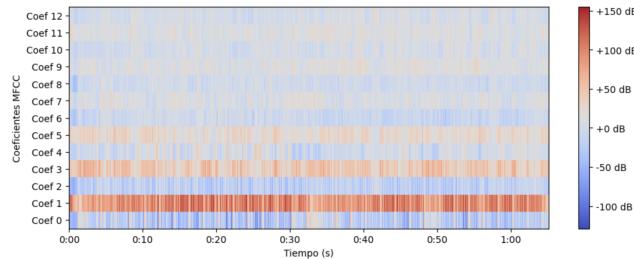
(b) The Beatles

Vemos que la amplitud de las ondas de BTR es, en general, más grande que las de los Beatles; esto se debe principalmente a los instrumentos utilizados en ambas canciones. En BTR se utilizan más los instrumentos eléctricos/digitales, en cambio que con los Beatles, se utilizan más los eléctricos/acústicos. También cabe mencionar que el género tocado por BTR es más "ruidoso" o "fuerte" que el de los Beatles.

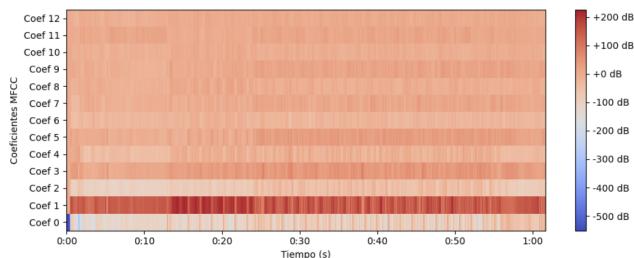
3. Descripción de Mel Frequency Cepstral Coefficients

Veamos ahora una representación gráfica de los coeficientes frecuencia. A partir de estos, podemos extraer características basadas en la percepción humana del sonido.

Figura 3: MFCC



(a) Big Time Rush



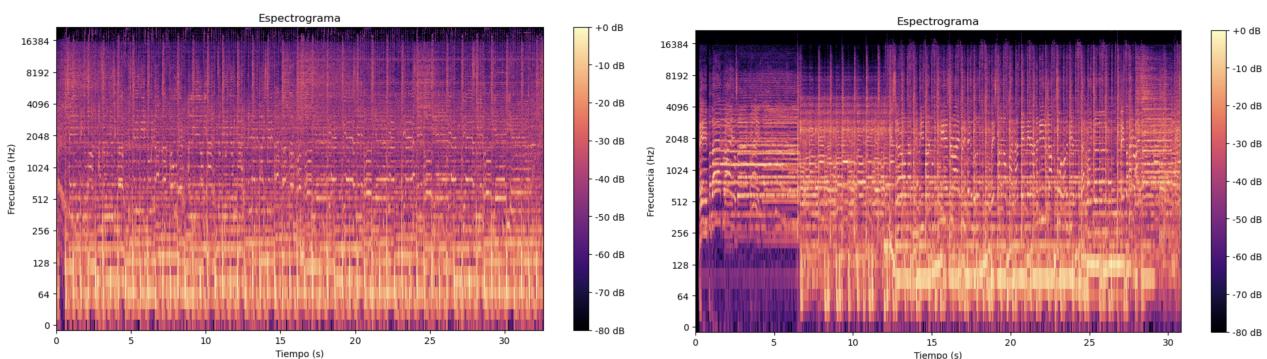
(b) The Beatles

Podemos notar que hay una diferencia muy notable en ambas pistas. La de BTR nos indica que el sonido con más potencia es percibido por el ser humano más en las frecuencias bajas y medias que en las altas, mientras que en la de los Beatles, nos indica que el sonido con más potencia es percibido por el ser humano más en las frecuencias altas que en las medias y bajas.

4. Descripción de espectogramas

Veamos ahora una representación visual de la intensidad de frecuencias a lo largo del tiempo.

Figura 4: Hercios x segundo



(a) Big Time Rush

(b) The Beatles

Viendo ahora el contenido auditivo desde otro punto de vista, vemos como se comportan las frecuencias y sus potencias a lo largo del tiempo.

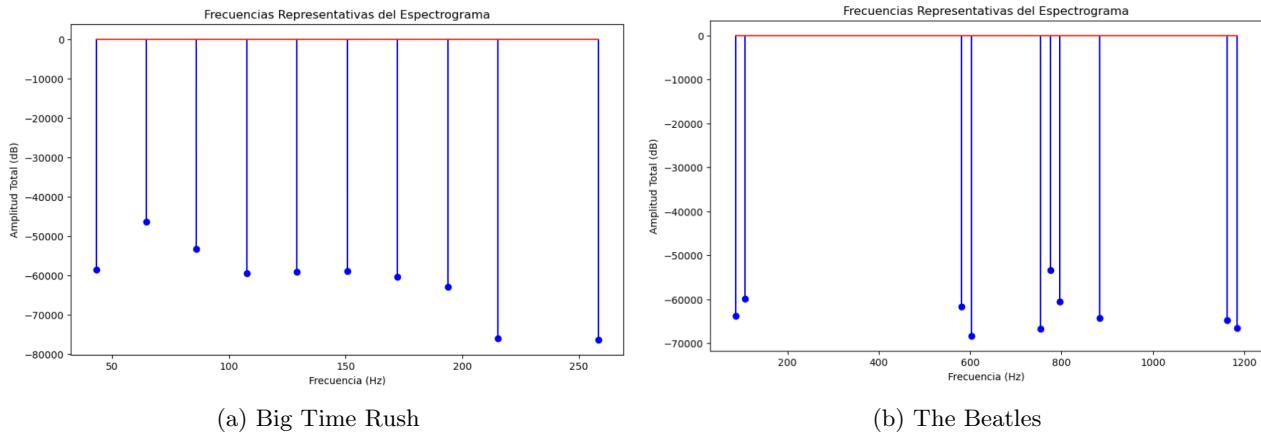
Big Time Rush mantiene su composición de manera uniforme en los 34 segundos analizados. Mayor presencia y potencia en las frecuencias graves que en las medias y agudas.

Los Beatles no mantienen su composición de manera uniforme en los 31 segundos analizados. En los primeros 7 segundos hay mayor presencia y potencia en las frecuencias medias y agudas, pero después la mayor presencia y potencia se inclina a las frecuencias graves.

Podemos ver, inclusive, que aunque Los Beatles no mantienen de manera uniforme la presencia y potencia de sus frecuencias en el segmento analizado, la potencia y presencia de sus frecuencias es más notoria en su canción que en la de BTR.

5. Significancia

Por último, analicemos únicamente el comportamiento de la amplitud de las frecuencias más representativas.



Big Time Rush. Las frecuencias más significativas van desde los 45 a 300 hercios. Y la mayoría de las amplitudes de sus ondas se encuentran entre los -45,000 y -75,000 decibeles. Sus frecuencias están dentro del rango grave.

The Beatles. Las frecuencias más significativas van desde los 100 a 1,200 hercios. Y la mayoría de las amplitudes de sus ondas se encuentran entre los -50,000 y -70,000 decibeles. Sus frecuencias están dentro del rango medio bajo y medio.

En resumen, la canción de los Beatles abarca frecuencias significativas más extensas. Pero la canción de Big Time Rush abarca amplitudes de onda significativas más extensas.

4. Conclusión

Hemos ahora introducido al análisis auditivo temas de algoritmos y probabilidad, y como estos nos pueden ayudar a tener una primera idea de lo que sucede detrás de la grabación de una canción o un audio en general. Fue muy interesante poder analizar a instrumentos y voces en conjunto, pero sería ahora más interesante poder hacer una comparativa de ellos por individual; sin mezclarlos.

Aun queda un largo camino por recorrer, pero lo importante es que estamos aprendiendo a realizar nuestros propios análisis. La aplicación de este modelo fue utilizada para el tema musical (canciones), pero estoy seguro que hay análisis todavía más profundos que nos pueden ayudar a entender como es que los dispositivos como google o Alex reconocen indicaciones de voz por parte de una fuente externa.

5. Referencias

- [1] Estefania Coluccio Leskow. Sonido. En concepto, una enciclopedia online más viva, más simple y más confiable, 2024. URL <https://concepto.de/sonido/>. Accedido el 16 de marzo de 2025.
- [2] Fanny Pirela Sojo. Música. En concepto, una enciclopedia online más viva, más simple y más confiable, 2024. URL <https://concepto.de/musica/>. Accedido el 16 de marzo de 2025.