



Tecnológico de Monterrey

Campus CCM

Análisis de Métodos Matemáticos para la Física (Gpo 631)

Dra. Marisol Rodriguez Arcos

Fase 2 Situación Problema

Alejandro Santiago Baca Eyssautier
César Daniel Yamada Xochipa
Yamilet Lozada Rangel
David Axel Rodríguez Botello
Ignacio Reza Vera

A01656580
A01658233
A01664113
A01663282
A01658364

INTRODUCCIÓN

La similitud entre canciones es un tema controversial en la industria musical. En un mundo tan globalizado, los artistas tienden a tomar inspiración de otras obras, caen en tendencias musicales en las que patrones sonoros se vuelven populares o simplemente recurren al plagio deliberado.

Con el fin de desarrollar una aplicación innovadora especializada en identificar la similitud entre canciones, se desarrollará un código que analizará fragmentos de canciones utilizando principios basados en análisis espectral de las composiciones musicales como lo hacen las plataformas de Shazam o Soundhound. Para hacer este análisis, estas aplicaciones utilizan algoritmos fundamentados en la transformada de Fourier, el cual es un concepto fundamental para el procesamiento de señales de audio ya que descompone estas señales en sus componentes individuales de frecuencia.

Esta aplicación persigue la capacidad de monitorear las emisiones de radio, tanto análogas como digitales, con el fin de asegurar el cumplimiento de los derechos de autor y el debido pago de regalías. Su propósito primordial radica en salvaguardar los intereses de los creadores musicales al asegurar que sus obras sean debidamente reconocidas y compensadas en todas las transmisiones públicas. Para lograr este cometido, se aspira a emplear tecnologías de vanguardia y conceptos matemáticos que garanticen la exactitud y fiabilidad en la identificación de obras musicales en diversas plataformas de emisión.

MARCO TEÓRICO

El sonido se propaga en forma de ondas a través de un medio, como el aire. Estas ondas se caracterizan por su amplitud, frecuencia y fase. Cada sonido puede descomponerse en un conjunto de frecuencias que el ser humano puede oír, generando un espectro de frecuencia. En la música se pueden encontrar sonidos asociados a frecuencias específicas a las que se les refiere como notas musicales, las cuales se organizan en escalas, siendo la más común la escala diatónica.

Nota	Do (C)	Re (D)	Mi (E)	Fa (F)	Sol (G)	La (A)	Si (B)
Rango de frecuencia	261.63 Hz	293.66 Hz	329.63 Hz	349.23 Hz	392.00 Hz	440.00 Hz	493.88 Hz

Tabla 1. Notas de la escala diatónica y sus rangos de frecuencia aproximados.

La música se puede interpretar como señales que recibimos con información que podemos representar de una manera más sencilla. La teoría de señales proporciona el fundamento para comprender y manipular señales, incluyendo conceptos como el muestreo, la discretización y la representación de señales en el dominio del tiempo y la frecuencia. Estas señales se manipulan mediante el proceso digital de señales (DSP) que se comprende de algoritmos y técnicas para el procesamiento de señales digitales, incluyendo filtros, técnicas de reducción de ruido, y la implementación de operaciones matemáticas en señales, consiguiendo así la altura del sonido o tono y el timbre, que permite distinguir diferentes instrumentos y voces.

Ahora bien, las señales también proporcionan el espectro de frecuencia que muestra elementos como la distribución de la energía o amplitud de las componentes de la frecuencia que se obtienen de esa señal, o dicho de otra forma, las notas musicales. Este espectro de frecuencia se representa utilizando la transformada de Fourier, herramienta matemática que proporciona la representación en el dominio de la frecuencia de la función que describe la señal. Dicha transformada de Fourier es continua para señales análogas y discreta en el caso de señales digitales.

Transformada de Fourier Continua (FT):

$$X(\omega) = \int_{-\infty}^{\infty} x(t) \cdot e^{i\omega t} dt$$

Donde:

- $X(f)$ es la representación en el dominio de la frecuencia de la señal continua
- $x(t)$ es la señal en el dominio del tiempo continuo
- ω es la frecuencia angular $\omega = 2\pi f = \frac{2\pi}{T}$
- i es la unidad imaginaria

Transformada de Fourier Discreta (DFT):

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-i \frac{kn}{N}}$$

Donde:

- $X(k)$ es la k-ésima componente de frecuencia de la señal en el dominio de la frecuencia.
- $x(n)$ es la muestra de la señal en el dominio de tiempo.
- N es el número total de muestras en la señal (longitud).
- i es la unidad imaginaria

Cuando se habla de análisis computacional, la eficiencia es muy valiosa y las transformadas llegan a ser muy costosas computacionalmente. No obstante, la transformada rápida de Fourier (FFT) permite reducir el número de operaciones necesarias para calcular la transformada de Fourier Discreta ya que reduce la complejidad al utilizar menos operaciones aritméticas. Esto se logra gracias al enfoque para realizar cálculos eficientes al dividir la secuencia en secuencias más cortas, ahorrando así tiempo de procesamiento. El algoritmo de esta transformada es el siguiente:

La FFT $y[k]$ de longitud N de la longitud $-N$ la secuencia $x[n]$ se define como

$$y(k) = \sum_{n=0}^{N-1} e^{-2\pi j \frac{kn}{N}} x[n]$$

y la transformada inversa se define de la siguiente manera

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} e^{2\pi j \frac{kn}{N}} y[k]$$

MÉTODOS

Como método de investigación decidimos comparar 2 canciones que se consideran parecidas. Para decidir objetivamente dos canciones similares consultamos listas de canciones con acusaciones de plagios y finalmente decidimos analizar “*Blame It on the Boogie*” de The Jacksons y “*Será que no me amas*” de Luis Miguel.

"Blame It on the Boogie" es una canción disco y funk interpretada por The Jacksons, lanzada en 1978 del álbum *Destiny*, la canción tiene una duración aproximada de 3 minutos y 36 segundos y la pista ha sido reconocida como un clásico de la música disco. Por otro lado "Será Que No Me Amas" es una canción romántica interpretada por Luis Miguel, lanzada en 1988 como parte del álbum "Busca Una Mujer", la canción tiene una duración aproximada de 4 minutos y 48 segundos. Es conocida por su emotividad y por ser una de las canciones emblemáticas del repertorio de Luis Miguel.

Para llevar a cabo esta comparativa, analizaremos 2 momentos claves:

1. Los primeros 10 segundos, que corresponden a la introducción de ambas canciones, para poder analizar púramente la melodía de estas.
 - a. Adicionalmente, se realizará una comparativa más precisa a partir del segundo 3 al segundo 6.
2. Del segundo 35 al 43, correspondientes al primer coro de las canciones, para analizar de igual forma el cómo afecta la adición de la voz al análisis espectral.

Dicho análisis espectral será llevado a cabo con el siguiente código:

```

# Importamos las librerías necesarias
import numpy as np
import scipy.io.wavfile as wav
import matplotlib.pyplot as plt
from scipy.signal import medfilt

def analisis_espectral(archivo_wav, inicio_segundo, fin_segundo, index, name):
    # Leer cierto fragmento del archivo de las canciones
    tasa_muestreo, datos = wav.read(archivo_wav)
    inicio_muestras = int(inicio_segundo * tasa_muestreo)
    fin_muestras = int(fin_segundo * tasa_muestreo)
    datos = datos[inicio_muestras:fin_muestras]

    # Aplicar un filtro de mediana para reducir el ruido en los datos
    datos_filtrados = medfilt(datos, kernel_size=3)

    # Obtener la magnitud de la Transformada de Fourier
    transformada_fourier = np.fft.fft(datos_filtrados, norm='forward')
    magnitud = np.abs(transformada_fourier)

    # Calcular las frecuencias correspondientes
    n = len(datos)
    frecuencias = np.linspace(0.0, tasa_muestreo / 2, n // 2)

    # Graficar el análisis espectral
    plt.subplot(1, 2, index+1)
    plt.plot(frecuencias, 2.0/n * magnitud[:n//2], color='deeppink')
    plt.title(f'{name}', fontstyle='italic')
    plt.xlabel('Frecuencia (Hz)', fontstyle='italic')
    plt.ylabel('Amplitud', fontstyle='italic')
    plt.grid(True, linestyle='--', alpha=0.6)

```

Introduciendo los valores y llamando a la función de la siguiente manera:

```

# Lista con el nombre de los archivos y título de las canciones
labels = [["Será Que No Me Amas.wav" , "The Jacksons Blame It On the Boogie.wav" ],
          ["Será Que No Me Amas - Luis Miguel", "Blame It On The Boogie - The Jacksons"]]

# Fragmento de ambas canciones por analizar
inicio_segundo, fin_segundo = 3, 6

# Graficar el análisis espectral de las canciones
plt.figure(figsize=(15, 5))
for i in range(len(labels)):
    analisis_espectral(labels[0][i], inicio_segundo, fin_segundo, i, labels[1][i])
plt.suptitle(f'Análisis Espectral de dos canciones del segundo {inicio_segundo} al {fin_segundo}', fontweight='bold')
plt.tight_layout()
plt.style.use('dark_background')
plt.show()

```

En este código se utilizó la función `.fft` de la librería de `numpy`, de la siguiente forma:

```
numpy.fft.fft(a, n=None, axis=-1, norm=None)
```

Donde:

- `a`: (array_like)
 - Parámetro principal en forma de arreglo que contiene la secuencia que se va a transformar.
- `n`: (int, optional)
 - Parámetro opcional que especifica el tamaño de la salida. Si no se especifica este parámetro, el tamaño del arreglo proporcionado (`a`).
- `axis`: (int, optional)
 - Indica el eje en donde se calcula la transformada.
- `norm`: ({“backward”, “ortho”, “forward”}, opcional)
 - Aplica la normalización para ajustar la amplitud de las componentes utilizando factores de escala. Backward es la opción predeterminada que no normaliza. Ortho aplica un factor de escala $1/\sqrt{N}$ que preserva la energía total. Forward aplica un factor de escala $1/N$ que preserva la energía en el tiempo y la frecuencia.

Es importante mencionar que en este proceso se normalizó el audio y se aplicó una técnica de reducción de ruido, para disminuir el ruido no deseado en las señales antes de realizar el análisis espectral. Esto último produciendo que la segunda ventana de frecuencia correspondiente a la fase de frecuencias, sea eliminada.

Lo anterior se llevó a cabo, debido a que, en muchos casos la fase no es relevante para el análisis general de la señal de audio y suele ignorarse en aplicaciones como el procesamiento de audio o análisis de música como es nuestro caso. Por lo cual, nos enfocamos únicamente en la magnitud de las frecuencias presentes en la señal de audio, ya que esta es la información que muestra qué frecuencias están presentes y la amplitud de cada una en esa sección de la canción.

RESULTADOS

- Para la introducción a ambas canciones, del segundo 1 al 10, se obtuvo lo siguiente:

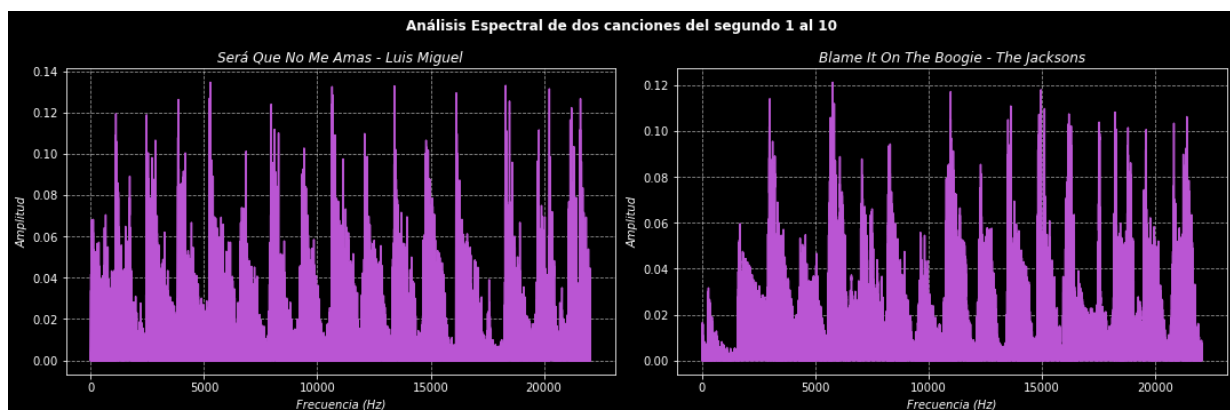


Figura 1: Análisis Espectral de los fragmentos de canciones “*Blame It on the Boogie*” de The Jacksons y “*Será que no me amas*” de Luis Miguel, del segundo 1 al 10.

- En particular, reduciendo el análisis de esta introducción, del segundo 3 al 6, se obtiene:

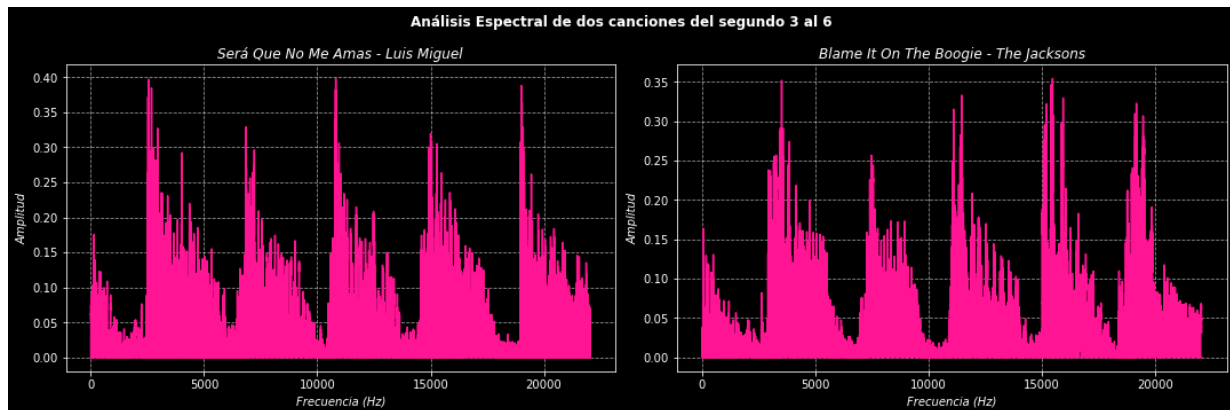


Figura 2: Análisis Espectral de los fragmentos de canciones “*Blame It on the Boogie*” de The Jacksons y “*Será que no me amas*” de Luis Miguel, del segundo 3 al 6.

- Finalmente, el análisis espectral del coro de ambas canciones, del segundo 35 al 43, se ve de la siguiente forma:

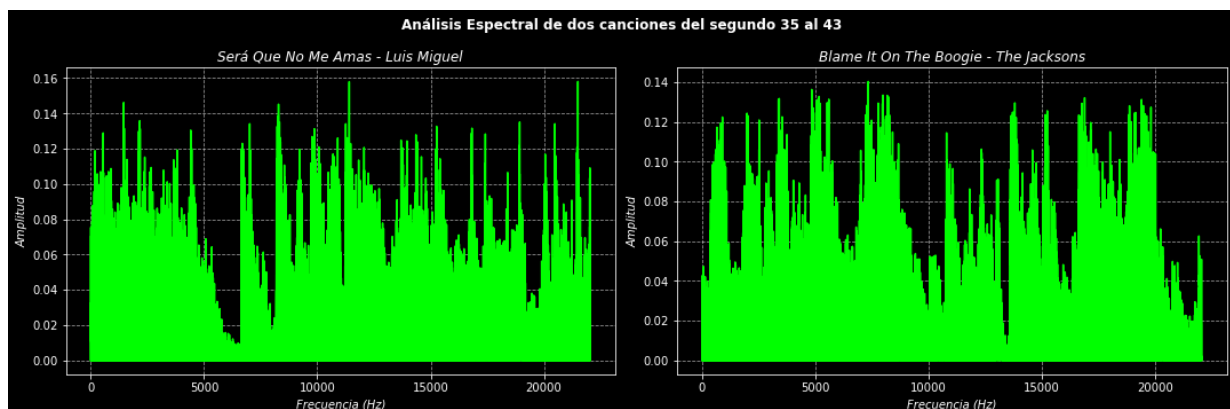


Figura 3: Análisis Espectral de los fragmentos de canciones “*Blame It on the Boogie*” de The Jacksons y “*Será que no me amas*” de Luis Miguel, del segundo 35 al 43.

ANÁLISIS DE RESULTADOS

Mediante las gráficas obtenidas, la Figura 2 ha demostrado ser la más reveladora en términos de similitudes y diferencias entre las canciones analizadas. Esta figura representa de manera exclusiva los componentes sonoros de las canciones analizadas, excluyendo las voces. Los resultados indican una notable similitud en las características espectrales indicando que pudiera haber cierta presencia de patrones idénticos. Estas observaciones respaldan la hipótesis de plagio en las composiciones.

Por otro lado, la Figura 3 presenta más limitaciones en términos de utilidad ya que esta incluye las voces en la representación espectral lo que ha dificultado la identificación precisa de patrones instrumentales, reduciendo así la capacidad de identificar las similitudes significativas. Sin embargo, es importante destacar que la presencia de voces no es suficiente para descartar la posibilidad de similitudes en otras capas de las composiciones.

La presencia de picos que coinciden en la Figura 2 pueden indicar la existencia de instrumentos o frecuencias específicas que se presentan de manera similar en ambas composiciones. Esto puede sugerir la utilización de instrumentos similares o la reproducción de patrones melódicos específicos. Igualmente pueden referir a la reproducción de notas o acordes que dan la percepción de similitud.

CONCLUSIONES

Al realizar las gráficas de comparación, por primera vez nos percatamos que diferían mucho entre sí, esto debido a la amplitud de cada canción; para solucionar esto se normalizó el audio, además se realizó un filtro de reducción de ruido, lo cual como se había mencionado previamente se hizo para concentrarnos exclusivamente en la magnitud de las frecuencias presentes en la señal de audio. De las gráficas nos damos cuenta que entre los segundos 3 y 6 de ambas canciones es donde más se parecen entre sí, esto puede deberse a la falta de voces en esas partes, ya que si comparamos con la figura 3, donde se comparan otros tiempos con la presencia de voces, nos damos cuenta que la similitud es menor. Por lo cual, si quisiéramos mejorar la comparación se deberían usar canciones con ausencia de voz, es decir canciones tipo acústicas.

De acuerdo al análisis espectral basado en la transformada de Fourier, podemos observar el considerable parecido entre las canciones, pero esto no es suficiente para determinar si es o no plagio musical, ya que este análisis se centra en la magnitud de las frecuencias presentes en una sección de la canción y no aborda otros aspectos cruciales del material musical, por lo cual no se puede concluir si hubo o no plagio entre las canciones. Para evaluar si hay plagio musical se requeriría un análisis más completo y detallado que involucre más aspectos de la composición musical.

REFERENCIAS:

Cetta, P. C. Análisis espectral aplicado a la música [en línea]. Material didáctico para el Área de Composición del Doctorado de Música. Universidad Católica Argentina. Facultad de Artes y Ciencias Musicales, 2009. Disponible en:
<https://repositorio.uca.edu.ar/handle/123456789/9191>

C, S. J. (2021, February 22). Frecuencias del sonido y armónicos musicales. Miguitarraelectrica; Miguitarraelectrica.
<https://miguitarraelectrica.com/frecuencias-sonidos-armonicos/>

DeChicchis, J. E. (2016). A Music Identification Algorithm which Utilizes the Fourier Analysis of Audio Signals.
<https://www.dechicchis.com/assets/Joseph-DeChicchis-Music-Identification.pdf>

Flores, M. (enero de 2016). Extensión de la transformada de tiempo corto de Fourier para sistemas no lineales tipo Volterra. Instituto Politécnico Nacional. Repositorio Dspace.
http://literatura.ciidiroaxaca.ipn.mx:8080/xmlui/bitstream/handle/LITER_CIIDIROAX/261/Flores%20Figueroa%2c%20G.%20M.%2c%202016.pdf?sequence=1&isAllowed=y

Iglesias, J. (2014). La Transformada de Fourier y la alteración del sonido.
<http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-Iglesias.pdf>

Luis Miguel. (1990). Será Que No Me Amas [Canción]. En *20 años*. Sunset Sound Studios

numpy.fft.fft — NumPy v1.26 Manual. (2022). Numpy.org.
<https://numpy.org/doc/stable/reference/generated/numpy.fft.fft.html>

Oppenheim, A. V., & Schafer, R. W. (2010). Discrete-Time Signal Processing. Pearson.

Oliphant, T. (2005). Numpy (1.26). Linux/Mac/Windows. NumPy Developers.

Parra, A. (2020). TOP 10: Canciones que son plagio y tuvieron éxito mundial. Diario El Ciudadano - de Costa Rica.
<https://laagendacr.com/top-10-canciones-que-son-plagio-y-tuvieron-exito-mundial/>

Rabiner, L. R., & Gold, B. (1975). Theory and Application of Digital Signal Processing. Prentice-Hall.

Smith, J. O. (2011). Spectral Audio Signal Processing. W3K Publishing.

The Jacksons. (1978). Blame It on the Boogie [Canción]. En *Destiny*. Epic Records