

ANÁLISIS DE LA ENVOLVENTE TÍMBRICA CON PYTHON

MARCELO F. GODOY ¹, PEDRO SOBRAL ²

Universidad Nacional de Tres de Febrero, Introducción a la acústica y psicoacústica, Buenos Aires, Argentina ^{1,2}

En los últimos años el lenguaje de programación “Python” se ha destacado en el campo del procesamiento digital de señales (P.D.S) por lo que muchos profesionales de dicha industria han optado por el mismo, impulsando a su vez el desarrollo de aplicaciones inherentes a la ingeniería en sonido, tanto en el campo de la ciencia de datos como la inteligencia artificial. La investigación realizada se centró en explicar las bases del P.D.S en Python para el análisis espectral y dinámico de sonidos complejos, a la vez que comentar las diversas aplicaciones del mismo.

1.INTRODUCCIÓN

El procesamiento digital de señales P.D.S es un concepto relativamente nuevo, teniendo en cuenta que hasta principios de la década de los 60 's la ciencia del audio se apoyaba casi en su totalidad en el formato analógico. Esto fue cambiando progresivamente hasta que el formato digital se estableció como predilecto. Junto con el enorme avance científico en materia de inteligencia artificial, acompañó un avance en la utilidad del procesamiento digital de señales, como lo vemos reflejado por ejemplo en programas de reconocimiento de voz.

El objetivo de este trabajo de investigación es el de realizar mediante el lenguaje python un programa que analice dinámica y espectralmente una nota tocada en un piano, y entregue el gráfico correspondiente a la envolvente tímbrica de la misma en la totalidad de su duración, para luego explicar la funcionalidad dicho análisis y su relevancia para ciertos avances científicos en materia de inteligencia artificial.

A lo largo del informe se desarrollarán conceptos indispensables para entender el enfoque del experimento realizado, dentro de los que se encuentran: el timbre, Muestreo y el Análisis de Fourier.

2.CONCEPTOS PSICOACÚSTICOS.

Dentro de las características subjetivas del sonido, se encuentra el concepto de “timbre”, que por definición entendemos como un conjunto de factores que permiten diferenciar un sonido de

otro[1], a estos factores se los denomina envolventes. En un sonido complejo se puede estudiar la envolvente dinámica, es decir, cómo fluctúa el sonido a lo largo del tiempo, y la envolvente espectral, que no es más que suma de los parciales de un sonido complejo.

2.1 CONCEPTOS DE P.D.S

El primer concepto de P.D.S necesario para el análisis y cómputo del sonido es la conversión análogo-digital. Este proceso cuenta con dos pasos: El muestreo, y la cuantificación, Fig.1. Definimos muestreo como la medición de la amplitud de onda en cierto intervalo de tiempo. Un factor importante este proceso es la frecuencia de muestreo, es decir la cantidad de veces por segundo que se mide la amplitud, Un valor conocido de este parámetro es 44100 hz, ya que es el valor de frecuencia de muestreo estándar para grabaciones. Mientras tanto la cuantificación toma esos valores de amplitud y los aloja en una matriz de números complejos. Esta conversión análogo-digital es crucial para el análisis del sonido, ya que la computadora no entiende que son las ondas sonoras analógicas y necesita un proceso para convertirlas en lenguaje binario.

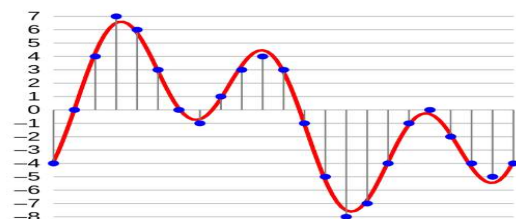


Fig.1. diagrama de muestreo de una señal

2.2 FAST FOURIER TRANSFORM (FFT) Y SHORT TIME FOURIER TRANSFORM (STFT).

En términos generales la transformada rápida de fourier (FFT) es un método matemático por el cual se puede descomponer un sonido complejo en una suma de sinusoidales (1), por lo que al graficarla es posible visualizar la envolvente espectral de dicho sonido ,fig.2.

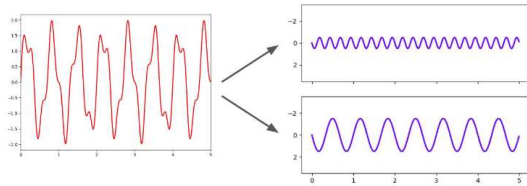


Fig.2. Descomposición de sonido complejo en suma de sinusoidales

$$s = A_1 \sin(2\pi f_1 t + \varphi_1) + A_2 \sin(2\pi f_2 t + \varphi_2) \quad (1)$$

La STFT cumple con la misma función que la FFT, con la diferencia que esta tiene en consideración la variable tiempo, es decir, analiza cómo fluctúa la suma de sinusoidales que componen el sonido a lo largo de un periodo de tiempo, por lo que se considera que conserva información del tiempo medido [2].

3. ENTORNO DE DESARROLLO

Python se caracteriza por ser un lenguaje de “alto nivel”, lo que quiere decir que su sintaxis es más parecida al idioma inglés que al lenguaje binario. Esto se da también porque cuenta con un intérprete incorporado. También cuenta con la posibilidad de importar “módulos” o “librerías”, los que amplían la ya amplia variedad de utilidades

3.1 MÓDULO “LIBROSA”

Librosa es una librería de código abierto centrada en el P.D.S con la cual se logra simplificar en gran medida el proceso de sampleo de cualquier señal, además de contar con una función específica para el cómputo de FFT y STFT.

3.2 MÓDULO “MATPLOTLIB”

Matplotlib es otra librería ampliamente utilizada en análisis de datos a nivel general, ya que cuenta con funciones que permiten generar distintos tipos de gráficos.

4. DESARROLLO DEL PROGRAMA

Para simplificar el desarrollo del programa se recurrió a una muestra de una nota tocada en un piano ya procesada y guardada en un archivo “wave form audio file format” (WAV), por lo que el proceso de conversión análogo-digital no fue parte en sí del programa, aun así este samplea en primera instancia la señal para generar una matriz donde aloja los con los datos que Luego usando la librería Matplotlib fueron graficados, Fig.3.

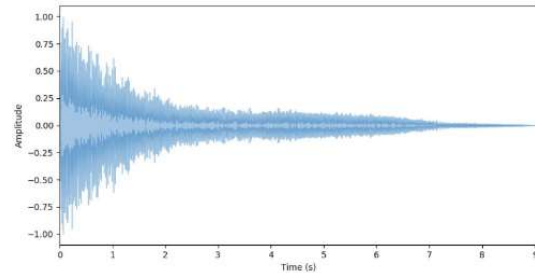


Fig.3. Gráfico amplitud vs tiempo de la nota de piano

El siguiente paso consistió en aplicar la función FFT del módulo librosa a la matriz donde se alojan los datos de la señal, que se representan con números reales. Luego de este proceso la matriz generadora del gráfico pasa un proceso de normalización lo cual permite graficar la transformada en coordenadas cartesianas, Fig.4, ya que sin este proceso el gráfico quedaría dado en coordenadas polares.

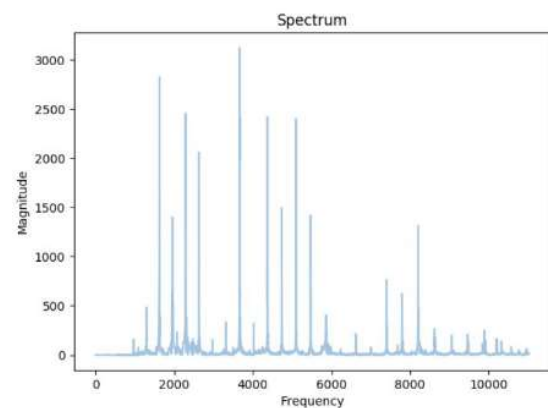


Fig.4. Gráfico de la FFT de la nota

La envolvente tímbrica es, en términos generales una combinación de la dinámica y la espectral, por lo que guarda información sobre frecuencia, amplitud, y tiempo, Fig.5. Por lo que solo la FFT no es suficiente para graficar dicha envolvente, ya carece de información del tipo temporal, es decir que esta es “estática”.

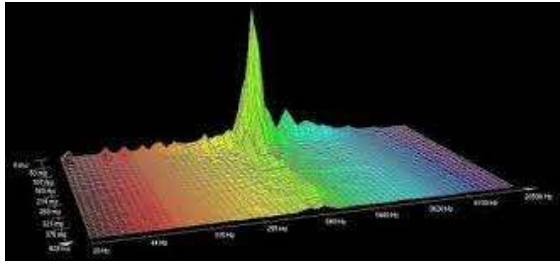


Fig.5. envolvente tímbrica

Finalmente el último proceso que lleva a cabo el programa es aplicar la función de la STFT, para esto es necesario establecer un número que indica la cantidad de transformadas rápidas de fourier que se quiere samplear por segundo, en este caso se utilizó una frecuencia de sampleo de 2048 hz, a su vez se estableció un periodo máximo de 9 segundos en la cual se ejecuta la función, si no fuera así el proceso sería infinito y resultaría en un “crasheo” del sistema.

5. RESULTADOS

Luego de que el programa realiza una análisis de la onda por medio de la STFT, nuevamente generamos un gráfico con la librería Matplotlib, en este caso un espectrograma, donde queda contenida información de la variación de amplitud y frecuencia en tiempo, de la onda, Fig.6. Es decir que mediante este proceso se consigue graficar la envolvente tímbrica en un periodo de tiempo finito.

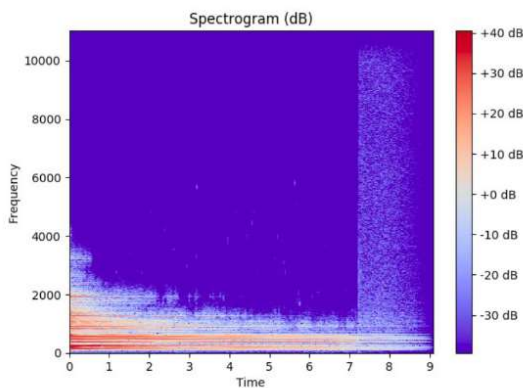


Fig.6. Espectrograma de la envolvente tímbrica de la onda

6. APLICACIONES

La funcionalidad del análisis realizado se hace notoria en el campo de la inteligencia artificial, un ejemplo de esto lo vemos en los programas de reconocimiento de voz que en primera instancia efectúa un análisis como el realizado por el programa desarrollado. La diferencia se encuentra en que el reconocimiento de voz en sí se da gracias a la aplicación de “redes neuronales”, Fig.7, que por medio de un algoritmo de inteligencia artificial asocian parámetros de la envolvente tímbrica con valores esperados por el sistema, pudiendo dar así una respuesta determinada.

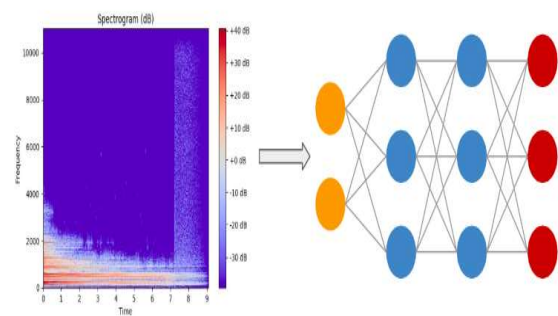


Fig.7. Gráfica de redes neuronales sobre espectrograma

7. CONCLUSIONES

Se puede concluir en que el procesamiento digital de señales está tomando una gran relevancia en la industria de las ciencias de datos y inteligencia artificial, esto gracias al gran desarrollo de herramientas para su utilización en este entorno que está teniendo lugar en el presente. Como ejemplo de esto se encuentran los repositorios de código abierto como “Github” y “Tensorflow” donde es posible encontrar algoritmos desarrollados exclusivamente para el P.D.S.

8. REFERENCIAS

- [1]A. I. Vistnes, Physics of Oscillations and Waves, Undergraduate Texts in Physics, ©Springer International Publishing Switzerland 2018
- [2]J. Unpingco Python for Signal Processing, ©Springer International Publishing Switzerland 2014

