

TP N°1 –Análisis de señales de audio

Este es un trabajo práctico inicial que introduce en forma pragmática algunos parámetros de análisis de señales acústicas. Se pensó además para revisar conceptos básicos de señales y simultáneamente algunos elementos de programación, como funciones, ciclos, vectorización y representación gráfica entre otras. El TP debe implementarse preferentemente en lenguaje Python.

El TP debe entregarse para su evaluación hasta el 10/10/22 inclusive. La entrega fuera de término será penalizada. Puede hacerse individualmente o en grupo de hasta tres personas. Sólo se debe entregar el código, en este caso no se pide informe. Si bien no se pide un análisis de los resultados en forma escrita, algunos puntos tienen asociados conceptos importantes del manejo de señales que son útiles y sobre los cuales es importante prestar atención. A partir de la guía pueden surgir muchas variantes que no necesariamente están contempladas en este TP y quedan abiertas a su curiosidad y exploración.

A su vez, se recomienda esforzarse en la claridad y organización del código para que facilite la lectura y análisis; incluir todos los comentarios que sean pertinentes. Cualquier duda o consulta sobre los incisos o el formato, pueden hacerlo en el foro del aula virtual correspondiente al TP1. La entrega deberán enviarla a la dirección pdsuntref@gmail.com.

1 Sintetizar una señal discreta, que simule una nota musical (LA 440), de 2 segundos de duración, muestreada a 44100 Hz. Considerar 5 componentes armónicas, la fundamental y cuatro armónicos consecutivos. Consideren una ley de decaimiento para la amplitud de los armónicos de $1/n$, en donde n es el número de armónico, siendo $n=1$ la fundamental de la señal. Normalizar a la señal final sintetizada $x[n]$ y graficar las componentes por separado y la señal total.

2 Para simular una señal de ruido, se les pide generar señales aleatorias, normalmente distribuidas (Gaussianas), con valor medio nulo y desviación estándar unitaria. Calcular para las longitudes (L) dadas en la tabla, la desviación estándar en cada caso y determinar la diferencia porcentual entre lo calculado y lo esperado ($\sigma = 1$). Utilizar funciones ya existentes (en las bibliotecas del lenguaje) para la generación de las señales aleatorias y crear su función propia para el cálculo del desvío estándar.

Valor medio: $\mu = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$

Desvío estándar: $\sigma = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (x[n] - \mu)^2}$ Varianza: σ^2

| L | σ | % |
|--------|----------|---|
| 5 | | |
| 10 | | |
| 100 | | |
| 1000 | | |
| 10000 | | |
| 100000 | | |

3 A la señal $x[n]$ sintetizada en el punto 1, sumarle ruido de media cero y desvíos de $\sigma = 0.1$, $\sigma = 1$ y $\sigma = 3.0$ para generar respectivamente tres nuevas señales $x01$, $x1$ y $x3$. Normalizar y graficar cada una.

Calcular la relación señal a ruido de las nuevas señales. La relación señal a ruido (SNR: *Signal to Noise Ratio*) se puede calcular con la siguiente definición [1] (hay muchas otras formas de calcularla, ver [2]):

$$SNR = \frac{\text{Amplitud}_{\text{señal}}}{\sigma_{\text{ruido}}}$$

Se suele utilizar un valor de $SNR = 3$ como límite de detección convencional, eso quiere decir que cualquier señal con una relación SNR superior a 3 se puede diferenciar del ruido. En caso que no se pueda medir el ruido por separado de la señal, se debe tratar de medir la desviación estándar del ruido en una zona donde se sepa, o presuma, que no hay señal, en el rango más amplio posible para contar con mayor estadística y mejorar la estimación. Analizar el efecto de agregar una componente de continua ($x[n] + c$) en la determinación de la SNR .

4 Generar 10 señales de ruido, diferentes entre sí, con media nula y $\sigma = 3$, y sumarlas a la señal del punto 1. Deben quedar 10 señales diferentes con la información de la señal más el ruido. Luego, calcular la señal promedio de todas las señales con el ruido agregado, calcular la relación SNR de la señal promediada y comparar el valor con la obtenida en el punto 3. Volver a repetir el punto con 100 y 1000 señales en el promedio. A esta técnica se la denomina promedio en el ensamble, o simplemente promediado.

Nota: si el generador de números aleatorios no tiene media cero se debe restar el valor medio antes de sumárselo a la señal.

5 Se implementará ahora un filtro de media móvil (*Moving Average, MA*) que responde a las ecuaciones de respuesta al impulso y relación entrada-salida que se dan a continuación. Esta es una de las tantas versiones posibles.

Se les pide implementar dos funciones, una directa que se llame $xfld = \text{mediamovild}(x, M)$, y otra implementación recursiva [3] $xfr = \text{mediamovildr}(x, M)$ que reciben como entrada a la señal x y la longitud de la ventana M del filtrado, y devuelve la señal filtrada $xfld$ o xfr . Comparar los tiempos de ejecución de cada una.

$$h[n] = \frac{1}{M+1} \sum_{k=0}^M \delta[n+k] \quad y[n] = \frac{1}{M+1} \sum_{k=0}^M x[n+k]$$

Una vez implementado, determinar algún valor particular de M , explicando el criterio utilizado, para dejar pasar la componente de 440 Hz de la señal del punto 1, atenuando al menos un 90% del resto de los armónicos. Graficar los resultados.

Nota de Precaución: la implementación de la referencia [3] del método recursivo usa otra forma del filtro (considera la ventana con muestras anteriores de la entrada).

Ayuda: para determinar el valor de M que filtre una determinada frecuencia se puede observar la cantidad de puntos que hay un periodo de la señal y en función de eso determinar el valor.

6 Otra forma de aplicar un filtro es mediante la convolución de la señal con la respuesta al impulso del filtro. Repetir el punto 5 pero realizando el filtrado mediante la operación de convolución lineal con la respuesta al impulso dada en el punto 5. Utilizar el mismo valor de M que se seleccionó en el punto 5. Comparar gráficamente las señales filtradas del punto 5 y 6, con amplitud normalizada.

7 Repetir el punto 6, pero la ventana a utilizar es ahora una ventana de Blackman (en lugar de la rectangular).

$$v[n] = a_0 - a_1 \cos\left(\frac{2\pi n}{M-1}\right) + a_2 \cos\left(\frac{4\pi n}{M-1}\right)$$
$$a_0 = 0.42 \quad a_1 = 0.5 \quad a_2 = 0.08$$

8 Convolucionar linealmente la señal sintetizada en el punto 1 con la respuesta al impulso *resp_imp.wav* (buscarla en la carpeta del aula virtual). Luego convolucionar en forma circular las dos señales con un largo igual a la señal respuesta al impulso. Finalmente, calcular con convolución circular, modificando las señales de forma tal de que el resultado sea igual a la convolución lineal. En todos los casos, graficar las señales convolucionadas y generar los archivos *.wav* correspondientes.

9 Implementar funciones individuales que permitan calcular los siguientes **parámetros temporales** de una señal discreta $x[n]$, con $0 \leq n \leq N-1$. Estas funciones deben estar disponibles para reutilizarse en otros algoritmos. Se seguirán las definiciones dadas en [4]. Aplicarlos a las tres señales *.wav* que figuran en la carpeta TP1 en el aula virtual y mostrar los resultados gráficamente.

- *Energía en tiempos cortos (Short Time Energy)*
- *Tasa de cruce por cero (Zero-Crossing Rate)*
- *Entropía de la Energía (Entropy of Energy)*

10 Implementar funciones individuales que permitan calcular los siguientes **parámetros frecuenciales** de una señal discreta $x[n]$, con $0 \leq n \leq N-1$. Estas funciones deben estar disponibles para reutilizarse en otros algoritmos. Se seguirán las definiciones dadas en [4]. Aplicarlos a las tres señales .wav que figuran en la carpeta TP1 en el aula virtual y mostrar los resultados gráficamente.

- *Centroide Espectral (Spectral Centroid)*
- *Flujo Espectral (Spectral Flux)*
- *Decaimiento Espectral (Spectral Rolloff)*

11 Graficar la magnitud transformada de Fourier (DFT) de las señales del punto 1 (sin ruido) y 3 (con ruido) utilizando tres ventanas diferentes: Rectangular, Hann y Blackman. La señal debe multiplicarse por la ventana antes del cálculo de la DFT. Además, graficar en escala de dB la DFT de cada una de las ventanas, y comparar el ancho del lóbulo principal y la atenuación.

12 Filtrar a la señal del inciso 3 con un filtro de media móvil, tal que deje pasar la componente de 880 Hz atenuada no más de -3 dB. Dar explícitamente el valor del ancho del filtro (valor de M) y graficar la magnitud de la respuesta en frecuencia del filtro y de la señal filtrada. Analizar si el filtrado es exitoso para todos los casos de ruido.

13 A la señal del inciso 1 calcularle y graficar la magnitud en dB y la fase de su STFT. Realizar tres gráficos cambiando el tipo de ventana y el largo del cálculo de la FFT. Determinar alguna combinación de las tres anteriores que sea (parezca) mejor para representar a la señal en estudio. Pueden utilizar una función “STFTS” estándar de alguna biblioteca (Ej. Scipy, LibRosa, etc).

Referencias:

[1] <https://terpconnect.umd.edu/~toh/spectrum/SignalsAndNoise.html>

[2] Steven W. Smith - The Scientist and Engineer's Guide to Digital Signal Processing - <http://www.dspguide.com/ch2.htm>.

[3] Steven W. Smith - The Scientist and Engineer's Guide to Digital Signal Processing - <http://www.dspguide.com/ch15.htm>.

[4] Eya Mezghani, Maha Charfeddine, Chokri Ben Amar, Henri Nicolas - Multifeature speech/music discrimination based on mid-term level statistics and supervised classifiers - 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA) - DOI: 10.1109/AICCSA.2016.7945728.