

Practica 1: Estudio de Audio con FFT

Univ. Juan Pablo Crespo Vargas

Resumen—En esta práctica se estudio el sonido y sus características, para lo cual se uso una computadora para grabar los sonidos, y el software MatLab para hacer el procesamiento digital del sonido.

I. RESUMEN TEÓRICO

Los sonidos son ondas muy complejas por tener muchos componentes de distinta índole que evolucionan rápidamente en el tiempo. Por lo cual su análisis y estudio consume mucho computo. Generalmente en acústica se trabajan con tres variables: El tiempo (t), la frecuencia (f) y la intensidad (I) o nivel sonoro.

Esto sugiere una representación gráfica tridimensional, que se representa en tres planos distintos: Plano dinámico (función del tiempo), plano espectral (función de la frecuencia) y plano melódico (frecuencia en función del tiempo) [1].

El espectro de sonido para el oído humano está entre las frecuencias de 20Hz a 20Khz, se conocen como infra sonido a las frecuencias por debajo a estas y ultra sonido a frecuencias mayores. Para el estudio del sonido y su relación con la música

La herramienta fundamental en el análisis espectral es la transformada rápida de fourier (fast fourier transform). Es un algoritmo que nos permite calcular en forma muy eficiente la DFT y su inversa, es el arma principal del manejo de señales digitales y filtros de este tipo. Con esta experiencia lo que se trata es de familiarizarse mas con el software de matlab, el cual nos brinda una gran cantidad de herramientas que de una u otra forma serán muy útiles.

I-A. Transformada Rápida de Fourier FFT

La transformada rápida de Fourier FFT es un algoritmo que reduce el tiempo de cálculo de n^2 pasos a $n \log_2(n)$. El único requisito es que el número de puntos en la serie tiene que ser una potencia de 2 (2^n puntos), por ejemplo 32, 1024, 4096, etc.

Las fórmulas con la que MATLAB calcula la transformada rápida de Fourier $Y = \text{fft}(x)$ y la transformada inversa $y = \text{ifft}(X)$ son, respectivamente:

$$X(k) = \sum_{j=1}^N x(j) w_N^{(j-1)(k-1)}$$

$$x(j) = \frac{1}{N} \sum_{k=1}^N X(k) w_N^{-(j-1)(k-1)}$$

La función fft de MATLAB convierte un vector de valores de la señal x en función del tiempo t en un vector g en función de la frecuencia w, $g = \text{fft}(x)$, g es un vector cuyos elementos son números complejos por que guarda información acerca

de la amplitud y de la fase. Ahora tenemos que asociar cada elemento del vector g con una frecuencia, del mismo modo que hemos asociado cada elemento del vector x con un tiempo.

El intervalo de frecuencias es $\delta w = \frac{2\pi fs}{n}$ de modo que la mínima frecuencia es 0 y la máxima $w_{max} = \frac{2\pi(n-1)}{n}$. La resolución espectral δw es inversamente proporcional al tiempo total $n\delta t$ de recogida de datos en la serie temporal.

I-B. Análisis musical

La nota La sirve como una referencia para todas las demás. A menudo se denomina nota de afinar. Se produce un La de afinar cuando el aire vibra a 440 veces por segundo, es decir a 440 hertzios. Por convención, a la octava que contiene esta nota La se le suele considerar la cuarta. Hay otra nota La, de una octava superior (la quinta octava) cuando el aire vibra a 880 hertzios, y otra más cuando vibra a $880 * 2$ (sexta octava), y otra a $800 * 2 * 2$ (séptima octava), etc, del mismo modo que hay un La que se produce cuando el aire vibra a $440/2$ (tercera octava) y otra a $440/4$ (segunda octava). Por convención, al La de la cuarta octava se la denomina La_4 , y análogamente al resto de las otras notas, con su nombre y el número de la octava en minúscula. Estas convenciones estandarizadas desde 1939 se conoce como el Índice acústico científico. Cada octava se extiende a través de un rango de frecuencias que es del doble del tamaño que el rango de frecuencias de la octava anterior. Para hallar la frecuencia de una nota cualquiera mediante una expresión matemática, se suele coger una frecuencia de referencia, por ejemplo el La de afinar (La_4 , 440 Hertzios) y se multiplica por la raíz duodécima de dos elevado al número de semitonos que separa el la de afinar de la nota que estamos buscando.

Con carácter general, una nota n ($n = 1$ para Do, $n = 2$ para Do#, $n = 12$ para Si) de la octava o (o desde 0 hasta 10) tiene una frecuencia $f(n, o)$ que podemos calcular mediante la ecuación 1.

$$f(n, o) = 440 * (2^{-12})^{(o-4)*12+(n-10)} \quad (1)$$

Para esta función, octava es un entero entre 0 y 10, y 'nota' es un entero en el rango de 1 a 12, de tal manera que $Do = 1$, $Do\# = 2$, $Re = 3$, $Re\# = 4$, $Mi = 5$, $Fa = 6$, $Fa\# = 7$, $Sol = 8$, $Sol\# = 9$, $La = 10$, $La\# = 11$, $Si = 12$.

II. DESARROLLO DEL TEMA

Para desarrollar los puntos, primeramente se muestran las características del equipo utilizado: Computador Intel Atom de 1.33Ghz, de 32 bits, operando MatLab2013 en Windows10.

II-A. Primera Etapa

Las premisas dadas en la guía son las siguientes:

1. Grabar señales de su propia voz y reconocer las frecuencias representativas, hasta el quinto, este será el timbre de voz principal.
2. Grabar notas musicales y construir sus espectros, seleccionar las cinco mas representativas.
3. Grabar notas de cuatro diapasones y encontrar la frecuencia característica y comparar.

Para la primera parte, en el programa 1 (anexo), se procede a grabar la voz durante 3 segundos, se leyó en un espacio con el menor ruido posible del uno al tres, en el gráfico 1 se muestra la captura en función del tiempo de los tres segundos, para este ejemplo se usaron los siguientes parámetros: Frecuencia de muestreo 8000, y 8 bits de resolución.

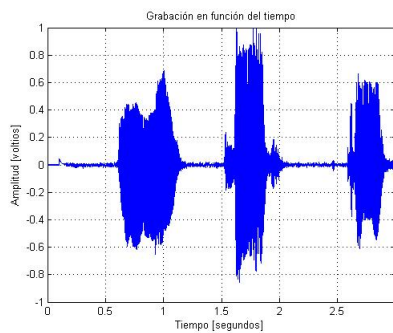


Figura 1. Grabación de 3 segundos de la voz

Una vez hecho el análisis de estos datos, mediante la transformada rápida de fourier, mostramos el periodograma de potencia vs. frecuencia en la figura 2.

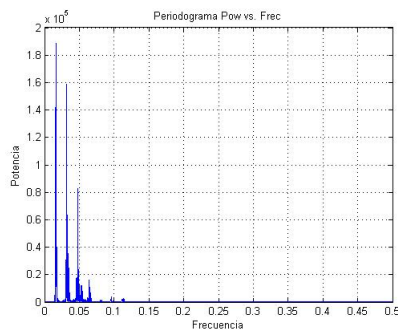


Figura 2. Periodograma

En la figura 3, se muestra a la potencia vs el Periodo, mostrando en esta figura el primer armónico, el principal.

Como siguiente punto, se muestran las gráficas 4 y 5 de diapasones de 440hz [2] y 528hz [3] respectivamente, haciendo el análisis de estos sonidos y obteniendo los timbres característicos de cada uno.

Como siguiente estudio, procedemos a grabar y analizar las notas principales de la tarka, en la figura 6 se muestran las notas de una tarka, se estudiarán las primeras 3; RE, MI y FA.

Las figuras 7, 8, 9 muestran los espectros de frecuencia de cada uno.

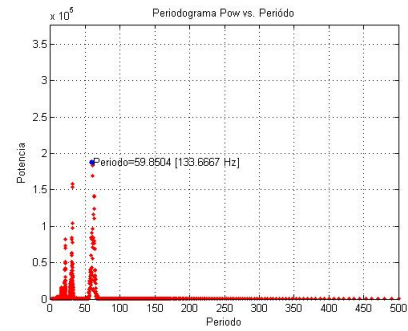


Figura 3. Periodograma

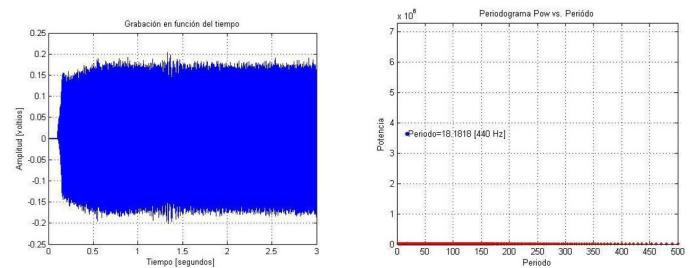


Figura 4. Representación en tiempo y espectro del diapason de 440 Hz

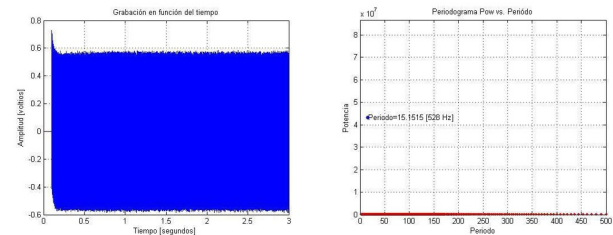


Figura 5. Representación en tiempo y espectro del diapason de 528 Hz

NOMENCLATURA DE LA TARKA

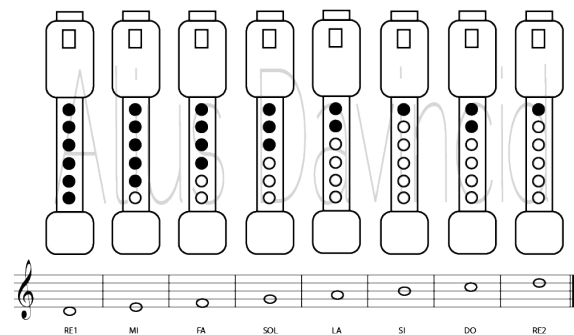


Figura 6. Notas de la Tarka [4]

II-B. Segunda Etapa

Convertir el espectro de Potencia Vs frecuencia, en un Sonograma. Consiste en reducir las tramas a uno de 32ms para las frecuencias y crear un nuevo eje de tiempo con lo

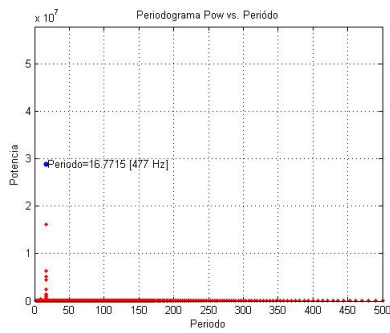


Figura 7. Notas de la Tarka [4] RE

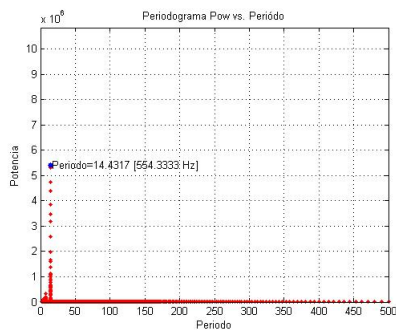


Figura 8. Notas de la Tarka [4] MI

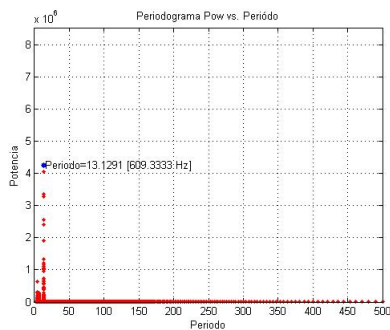


Figura 9. Notas de la Tarka [4] FA

que se formaran tres vectores y corresponderían al gráfico en 3D, sin embargo para el Sonograma se debe reducir en 2D con ayuda de colores o intensidad de grises o grosor. Como se describió, en la trama TT, hay un espectro de frecuencia, el siguiente TT otro y así, nTT donde n puede llegar a cubrir minutos u horas. Entonces, en un gráfico 3D en el eje Z estará la potencia, en Y la frecuencia y en X el tiempo TT. Esto puede simplificarse seleccionando colores que signifiquen la potencia en Y para cada tiempo - X convirtiendo en gráfico 2D.

Las premisas dadas en la guía son las siguientes:

1. Construir el programa que grabe tramas de 32ms, que use FFT para el espectro y ploteo en 3D y re grafique durante 2 minutos, en tiempo real.
2. Idéntico al anterior pero para gráfico 2D que es un Sonograma a colores, durante 2 minutos, en tiempo real.

3. Grabar notas de diapasones y encontrar la frecuencia característica y comparar.

Mostramos en las gráficas 10 y 11 la representación del sonograma de una nota La, para el diapasón 440Hz [2], ambas figuras muestran la frecuencia versus el tiempo, con la intensidad de la frecuencia como una intensidad de colores y escala de grises respectivamente.

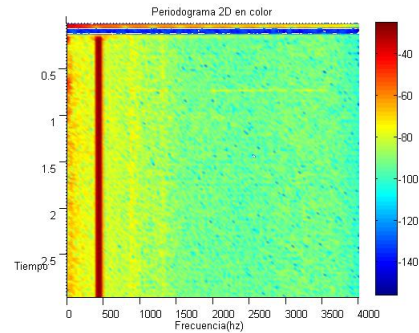


Figura 10. Sonograma del diapasón 440Hz a color

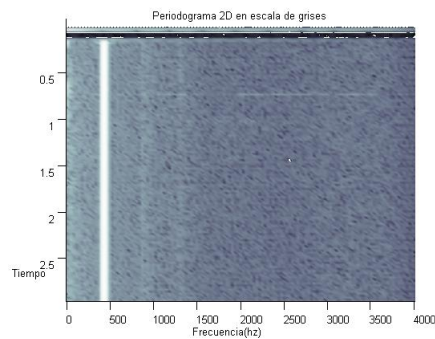


Figura 11. Sonograma del diapasón 440Hz a escala de grises

Mientras que en la figura 12 se muestra la representación tridimensional del objeto acústico.

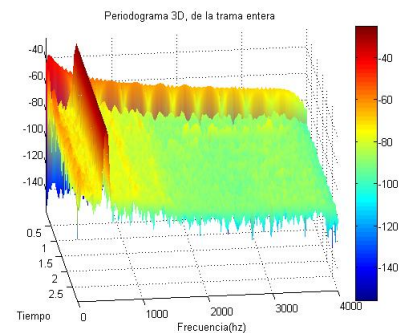


Figura 12. Sonograma del diapasón 440Hz en 3D

De igual manera las gráficas 13 y 14, 15 y 16, 18 y ?? muestran los sonogramas en 2D de las notas RE, MI y FA de la tarka, mismas que analizamos en la anterior sub sección.

De la misma forma mostramos los sonogramas en 3D en las figuras 19, 20 y 21.

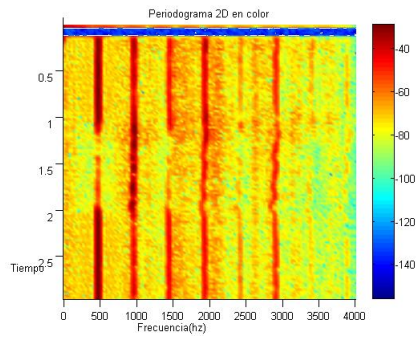


Figura 13. Sonograma 2D de la nota RE

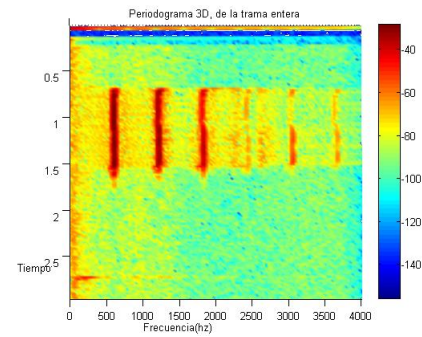


Figura 17. Sonograma 2D de la nota FA

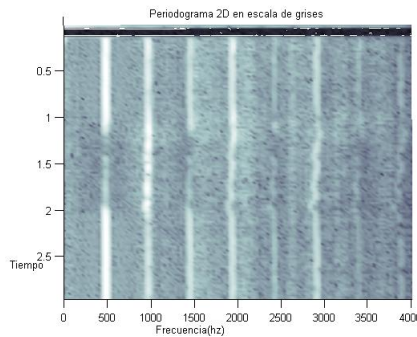


Figura 14. Sonograma 2D de la nota RE

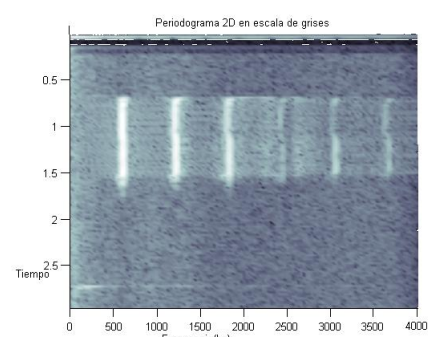


Figura 18. Sonograma 2D de la nota FA

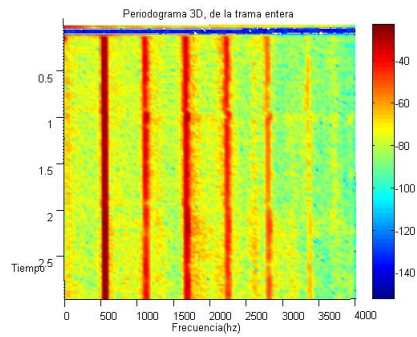


Figura 15. Sonograma 2D de la nota MI

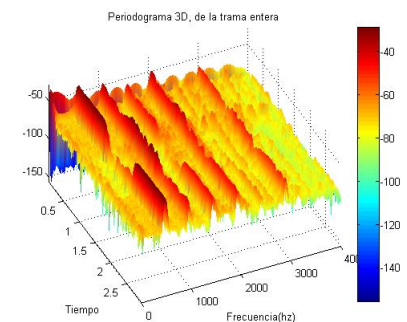


Figura 19. Sonograma 2D de la nota RE

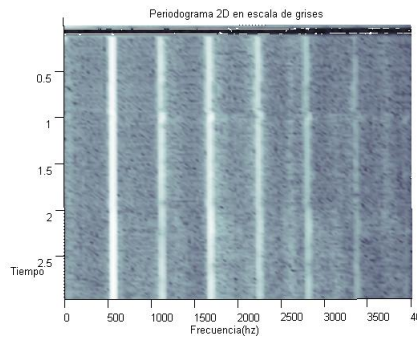


Figura 16. Sonograma 2D de la nota MI

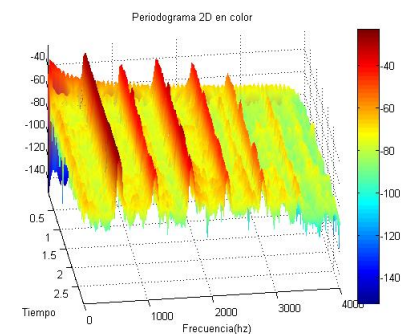


Figura 20. Sonograma 2D de la nota MI

II-C. Aplicación

Para la generación o reconocimiento de sonidos, debemos disponer un banco de datos en frecuencia de al menos hasta el

sexto armónico de tonos y semitonos, estos formaran el timbre de voz o el de un instrumento específico. Las frecuencias

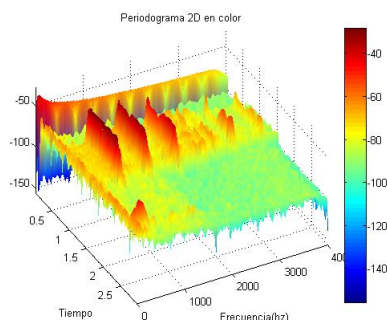


Figura 21. Sonograma 2D de la nota FA

principales o primer armónico corresponden a las notas del sonido. Con estos sonidos pueden formarse cadencia de notas musicales o reconocimiento de notas musicales para construir un pentagrama o iniciarse en el reconocimiento de la voz, todo esto mediante el método de comparación. En Internet se pueden encontrar diversas tablas para cada instrumento, tramas características para cada nota musical y lo que necesitamos las frecuencias típicas con seis armónicos, estas pueden ser descargadas, contienen para el piano los tonos y semitonos (acordes sostenidos o bemoles). En esta práctica usamos la ecuación 1 para generar todo el índice musical.

Se hace una comparación de las frecuencias con la frecuencia del primer armónico, si esta coincide con el valor con un error del 10 % se toma la nota. Sin embargo cabe notar que en el programa de redondea a la frecuencia entera mas cercana para hacer esta comparación.

Los datos al hacer correr el programa con la frecuencias de prueba diapason 440hz reconocen la octava y la nota: "octava=4
nota=LA"

III. ANÁLISIS DE RESULTADOS

III-A. Análisis de la voz

En las figuras 1 y ?? se muestra el análisis de la voz del autor de este texto, para el cual se vocalizo los números uno, dos y tres. En la figura 1 se puede ver la amplitud en función del tiempo, claramente se puede ver la vocalización de cada número. En la figura ??, el programa hace el análisis de las componentes de las cinco armónicas predominantes, lo que se conoce como característica y timbre particular de la voz, dando los siguientes datos: 133,667, 130,148, 135,123, 135,111, 201,111.

Estas frecuencias serian el timbre que es característico de mi voz, el cual es muy difícil coincidir con otras personas.

III-B. Análisis de notas

En las figuras ?? y ?? se analizó las notas dadas por un diapason. Estas notas fueron caracterizadas de manera correcta, pues se reconoció como primer armónico las notas pertenecientes a estos diapasones, algunas variaciones se dieron por no contar con un estudio para hacer las grabaciones en silencio total, es decir, para estas pruebas se tuvo mucho ruido de fondo.

Las notas de las tarkas analizadas fueron mucho mas interesantes al analizar, pues se identificó mediante las figuras 7, 8 y 9 las frecuencias características del instrumento, teniendo el valor principal de: 447Hz para RE, 554.33Hz para MI y 609.33Hz para Fa.

Al momento de analizar los sonogramas de estas notas en las figuras 19, 20 y 21, en su representación 3D podemos notar como estas notas, y en general, los sonidos de las Tarkas son bastante característicos, pues tienen un redoble en las frecuencias, es decir, que la nota vibra entre frecuencias con topes superiores e inferiores, pero mantiene su componente armónica en un valor cercano a la nota que se interpreta. Si comparamos las figuras de los sonogramas tridimensionales de la tarka, con por ejemplo, las de un diapason en 440Hz de la figura 12 notamos la rápida y compleja evolución respecto al tiempo de las ondas.

IV. CONCLUSIONES

En esta memoria de trabajo, se trató de resumir el análisis matemático y físico que se obtiene al estudiar las ondas sonoras. Para lo cual se tomo como objeto de estudio la voz humana, el sonido de diapasones y algunas notas de una tarka, instrumento Boliviano de viento.

Se utilizó una computadora para poder hacer la recopilación y digitalización de las ondas sonoras, utilizando el mismo hardware que se encuentra en un ordenador convencional, el micrófono interno y el software adecuado para analizarlo, en este caso MatLab.

Se mostró que a través de la potente herramienta del análisis por descomposición de armónicos, dado por la transformada de Fourier y haciendo uso de la visualización de datos, el estudio de los sonogramas como la amplitud en función del tiempo y los espectros en función del tiempo y amplitud se puede hacer una caracterización eficiente de los sonidos. Se realizó el estudio y caracterización de tres notas de la Tarka, y se mostraron sus respectivos sonogramas, así como su representación en función del tiempo, mostrando lo complejos que son estos sonidos característicos de este instrumento de viento.

También se analizó la recopilación y comparación de una frecuencia de un diapason, con notas del índice de frecuencias, cuya tabla se fabricó a partir del análisis matemático y musical. Gracias a esta comparación, y con un error del 10 % se pudo encontrar e identificar la nota analizada.

REFERENCIAS

- [1] EL SONAGRAMA: UNA REPRESENTACIÓN PRÁCTICA DE LOS SONIDOS Arnaud Gérard A. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1562-38232011000200005
- [2] Nota La diapason a 440Hz https://www.youtube.com/watch?v=3tDv_9kfyDM
- [3] Nota diapason a 528Hz <https://www.youtube.com/watch?v=5gRjSudwJao>
- [4] INSTRUMENTOS AUTÓCTONOS DE VIENTOS FAMILIA TARKAS <http://boliviaautoctona.blogspot.com/2012/06/tarkas.html>
- [5] El MatLab y su help que contiene explicación precisa.
- [6] <https://www.youtube.com/watch?v=stxtqkZzJ-U>