

Aproximación a la síntesis de la música a través del análisis de fourier

An approach to the synthesis of music through Fourier Analysis

Jimmy Alexander Cortes Osorio^{1*}, Andrew M. Knott², José Andrés Chaves Osorio³

Departamento de Física, Universidad Tecnológica de Pereira, Pereira, Colombia

*Corre-e: jacoper@utp.edu.co

Resumen— Se presenta una corta introducción a la síntesis de la música mediante el uso de la transformada de Fourier asistida por computadora mediante Matlab®.

Palabras clave— aula, Fourier, música, síntesis, transformada.

Abstract— A short introduction to music synthesis using the Fourier transform with computer assistance using Matlab® is presented.

Keywords— classroom, Fourier, music, synthesis, transform.

I. INTRODUCCIÓN

La música resulta ser básicamente una combinación coherente de sonidos y silencios que produce efectos psicológicos a quien la escucha; que interesa profundamente a los seres humanos y es transversal a casi todos los públicos. A través del análisis de principios fundamentales de la música como melodía, armonía y ritmo; se pretende mostrar (usando la experiencia en clase de los autores) la importancia de la transformada de Fourier (FT) como base para entender la música usando modelos físico matemáticos.

La FT (Fourier Transform) permite hacer la descomposición en senos y cosenos de las diferentes notas de variados instrumentos musicales tales como la flauta, el saxofón y el piano, de tal forma que se aprecien sus componentes de frecuencia, amplitud y fase para su posterior reconstrucción por síntesis aditiva. De igual forma, **se pueden evaluar algunas características básicas del sonido asociadas a su forma de onda tales como el tono y el timbre.**

II. CONTENIDO

A. El sonido [1][2]

Se puede definir como el conjunto de ondas mecánicas de presión longitudinal que viajan usualmente a través del aire, o

cualquier otro medio sólido, líquido o gaseoso, produciendo una perturbación que es captada por los oídos.

En la definición del sonido también debe considerarse tanto el fenómeno físico como el psicoacústico, ya que bajo la ausencia de un oyente puede existir un evento sonoro, pero no el evento auditivo.

La definición del ruido básicamente se adhiere al mismo concepto del sonido, pero se ajusta más al aspecto de un evento psicoacústico que es desagradable a la percepción humana. En general se puede expresar como un sonido no deseado.

1. La Intensidad [1][2]

Es la cantidad de energía acústica que contiene un sonido. La intensidad viene definida por la potencia, que a su vez está determinada por la amplitud de la onda sonora, permitiendo distinguir si el sonido es fuerte o débil.

Esta cualidad se mide con el sonómetro y típicamente los resultados se expresan en decibeles (dB). Una definición algo más ajustada indica que la intensidad está relacionada con la energía que atraviesa en un segundo, a la unidad de superficie perpendicular respecto a la dirección de la propagación de la onda sonora, por lo que se plantea que dicha intensidad sonora se mide en W/m^2 .

El concepto de la intensidad del sonido no es fácil de cuantificar, ya que el oído humano no le reconoce de manera lineal; es decir un sonido con una intensidad doble, no causa el doble de la sensación audible esperada.

2. La duración [1][2]

Es la cualidad que determina el tiempo de vibración de un objeto; por ejemplo, se pueden escuchar sonidos largos, cortos, muy cortos, etc. En la Figura 1 se ilustra la forma de onda y la duración en segundos del sonido de una flauta, un saxofón y una señal pura seno para una nota musical Do central (261 Hz).

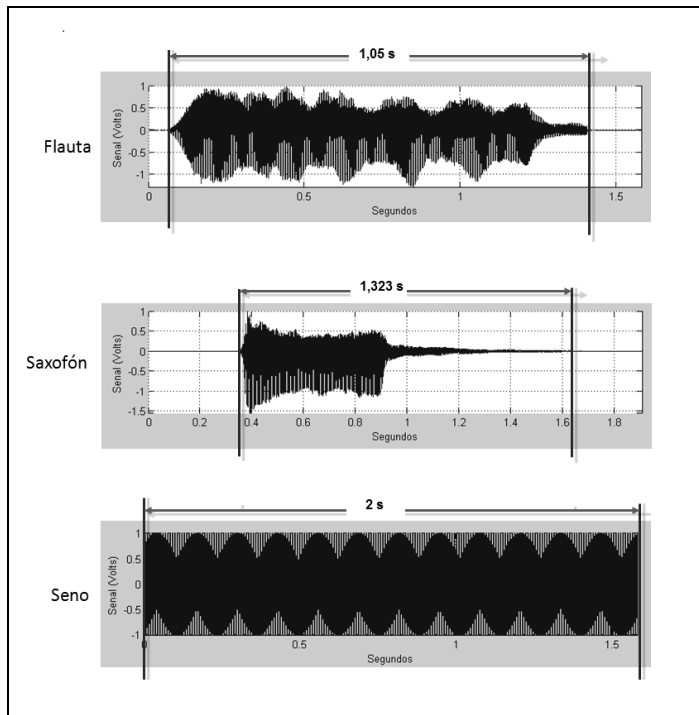


Figura 1. Forma de onda en el tiempo del sonido de una flauta, un saxofón y una función seno pura con su duración estimada.

3. El tono [1][2]

Éste viene determinado por la **frecuencia fundamental de las ondas sonoras medida en ciclos por segundo o hertz (Hz)**. El tono es quien **permite distinguir entre sonidos graves, agudos o medios**. Para que los humanos puedan percibir un sonido, éste debe estar comprendido en un rango entre **20 y 20.000 Hz** que se define como el **rango de frecuencia audible**.

Por debajo de este rango se tienen los infrasonidos y por encima los ultrasonidos. Es de observar que el rango de frecuencia audible varía ligeramente entre individuos, y para un mismo individuo, en la medida que avanza la edad el rango va reduciéndose tanto en graves como en agudos.

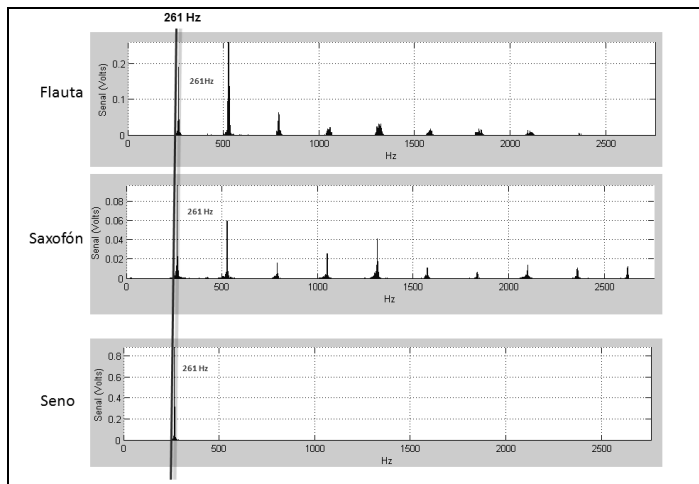


Figura 3. Análisis espectral de la flauta, el saxofón y la función seno pura.

En la Figura 3 se muestra el espectro en frecuencia de la nota Do central (261 Hz) para varios instrumentos musicales, tales como la flauta, el saxofón y una señal seno pura. Esta nota recibe el nombre de Do central, puesto que se encuentra muy cerca de la mitad del teclado del piano.

4. El timbre [1][2]

Es la cualidad que confiere al sonido el contenido de armónicos que acompañan a la frecuencia fundamental. El timbre caracteriza a cada instrumento haciéndole diferente aún bajo la misma frecuencia fundamental; esta cualidad es la que permite distinguir dos sonidos, por ejemplo, entre la misma nota musical (tono) con igual intensidad producida por dos instrumentos musicales distintos. **En los instrumentos musicales reales las frecuencias adicionales que componen un tono se denominan parciales, ya que la palabra armónico solo se ajusta a los múltiplos exactos de la frecuencia fundamental.**

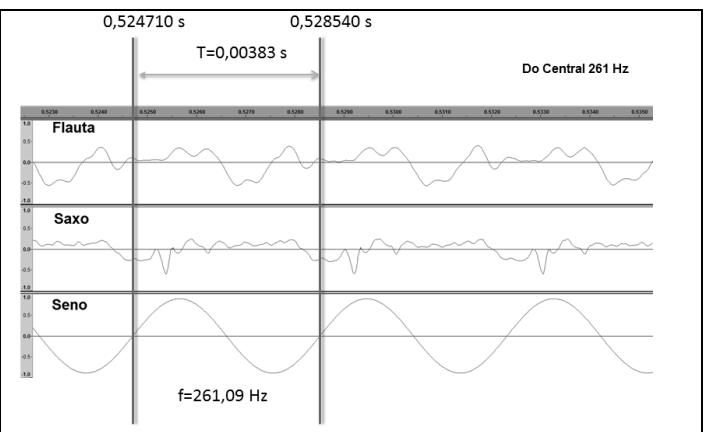


Figura 2. Forma de onda característica en el sostenido de la flauta, el saxofón y la función seno pura para un Do central.

En la Figura 2 se muestra que, aunque la nota Do central (de frecuencia de 261,09 Hz) **puede tener el mismo tono para diferentes instrumentos, el timbre es quien diferencia cada instrumento musical** (Nótese la forma de onda). El valor medido no es exacto y sólo se aproxima al verdadero de 261 Hz por definición.

B. El Teorema de Fourier [3][4]

Éste expresa que toda función periódica de frecuencia ω_0 puede expresarse como la suma infinita de funciones seno o coseno que son múltiplos enteros n de ω_0 . Se denomina ω_0 a la frecuencia fundamental y a cada término seno o coseno se le conoce como armónica. La ecuación 1 muestra su representación como una sumatoria:

$$f(t) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(n \omega_0 t) + (b_n \sen n \omega_0 t)]$$

Ecuación 1. Serie básica de Fourier.

La ecuación 1 también se puede expresar únicamente como una suma de funciones coseno de diferentes fases Φ_n y amplitudes A_n donde esta última es la magnitud de los coeficientes a_n y b_n (Véase la ecuación 2).

$$f(t) = a_0 + \sum_{n=1}^{\infty} [A_n \cos(n \omega_0 t + \phi_n)]$$

Ecuación 2. Serie de Fourier en función de cosenos.

Una representación más simplificada se logra mediante las series complejas de Fourier a través de la exponencial de Euler como se aprecia en la ecuación 3.

$$f(t) = c_0 + \sum_{n=1}^{\infty} [c_n e^{(n \omega_0 t)j} + (c_{-n} e^{-(n \omega_0 t)j})]$$

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{(n \omega_0 t)j}$$

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-(n \omega_0 t)j} dt$$

Ecuación 3. Serie compleja de Fourier y sus coeficientes.

1. Transformada discreta de Fourier [3][4][5]

Análogo a la serie, la transformada de Fourier descompone una señal periódica en senos y cosenos de diferentes frecuencias y amplitudes.

En las aplicaciones de ingeniería y tratamiento de señales, resulta más práctico considerar el proceso de manera discreta y no continua, ya que los sistemas de adquisición de datos no pueden obtener ni analizar la totalidad de la información.

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

Ecuación 4. Transformada Discreta de Fourier DFT

En la ecuación 4 se debe definir a:

N = Número de Muestras

n = Enésima muestra original

k = késimo termino de la DTF

Como alternativa a la Transformada Discreta de Fourier (DFT), se implementa el algoritmo de la Transformada rápida de Fourier (FFT), el cual realiza los mismos cálculos que la DFT, pero de manera mucho más rápida gracias a su recursividad. La ecuación 4 corresponde a la expresión que permite determinar la DFT.

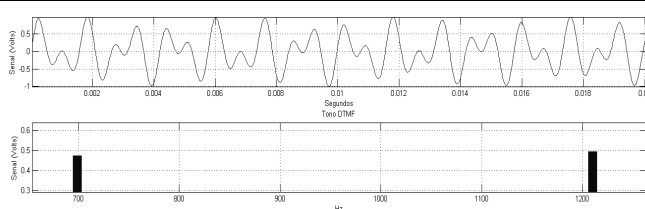


Figura 4. Señal DTMF en tiempo y en frecuencia

Una aplicación simple de la FFT, permite identificar los tonos y las amplitudes de las componentes de un tono Dual-tone multi-frequency signaling (DTMF) usado comúnmente en telefonía y otros medios de comunicación. Este teclado está compuesto por las frecuencias dadas en la figura 5:

Frecuencias de las Columnas				
	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Figura 5. Teclado telefónico típico con sus frecuencias DTMF correspondientes.

2. El sintetizador

Un sintetizador compone un nuevo sonido a partir de la combinación de elementos más simples, que de otra manera sólo existen en el mundo físico.

Los sintetizadores virtuales son los programas de computador capaces de reproducir el sonido de instrumentos musicales en tiempo real a partir de expresiones matemáticas cada vez más complejas dependiendo del realismo que se desea impregnar al sonido del instrumento sintetizado.

3. Síntesis Aditiva [6]

De manera clara, se puede inferir que la forma matemática más sencilla de realizar la síntesis aditiva es utilizando las **Series de Fourier, que permiten generar múltiples formas de onda modificando la amplitud y la fase de los parciales (armónicos) adicionados**. Ésta estrategia presenta como problema principal su alto consumo computacional al requerir un número importante de parciales, pero es muy apropiada a la hora de iniciarse en la construcción de una determinada señal.

El análisis para la Síntesis aditiva

Los autores han tomado como referencia la nota DO central o C4, la cual tiene una frecuencia fundamental cercana a 261Hz y se encuentra en la mitad del teclado del piano.

La mayoría de las interpretaciones musicales que ejecutan los aprendices de piano se encuentran alrededor de esta nota ya que es la más fácil de reconocer en la partitura (Véase la Figura 6).

Sib (La#)	466,16	Si	493,88
Lab (Sol#)	415,30	La	440,00
Solb (Fa#)	369,99	Sol	392,00
		Fa	349,23
		Mi	329,63
Mib (Re#)	311,13	Re	293,66
Reb (Do#)	277,18	Do	261,63

Figura 6. Figura del teclado con las notas básicas y sus respectivas frecuencias en Hz.

Para efectos de análisis se procede a utilizar el sonido instrumental de una **flauta y un saxofón**, previamente grabados en formato .wav, estas señales se comparan con respecto a una señal fundamental senoidal pura de la misma frecuencia.

Los dos primeros sonidos se capturan mediante un micrófono.

En la figura 1 se aprecian las señales obtenidas como funciones de tiempo, cada una de las cuales posee una determinada duración.

La figura 3 muestra el análisis espectral de estos instrumentos, apreciándose que todos ellos poseen como frecuencia fundamental 261 Hz, valor que es independiente de su timbre característico (Véase la figura 2).

En la Figura 7 se entrega el código en Matlab® que ha permitido la captura del sonido a partir de un archivo de audio *C-flauta.wav* previamente grabado por los autores. La gráfica generada se muestra en la Figura 1.

```

clc
clear all
%lee los datos del archivo de audio
[datos_leidos, fm, numbits]= wavread('C-flauta.wav');
muestras=length(datos_leidos);
%Se normaliza la senal sobre +/-1
datos_leidos=datos_leidos/max(datos_leidos);
%Reproduce la senal guardada a la frecuencia deseada
wavplay(datos_leidos, fm)
%Grafica la senal en tiempo
subplot(3,1,1)
duracion=muestras/fm;
tm=1/fm;
tiempo=linspace(0,duracion,muestras);
plot(tiempo, datos_leidos)
xlabel('Segundos')
ylabel('Senal (Volts)')
axis tight
grid

```

Figura 7. Código en Matlab® para la captura de los datos.

De otro lado, la Figura 8 ilustra el procedimiento detallado para el análisis básico de frecuencia y fase de la señal de audio del sonido *C-flauta.wav*.

```

%Grafica el espectro en frecuencia
subplot(3,1,2)
frecuencias=linspace(0, fm, muestras);
longmed=muestras/2;
fr=abs(fft(datos_leidos))/longmed;
bar(frecuencias, fr)
axis([0 fm/2 0 0.2])
xlabel('Hz')
ylabel('Senal (Volts)')
grid

%Grafica la fase
angulo=(angle(fft(datos_leidos)));
subplot(3,1,3)
plot(frecuencias, angulo)
axis([0, fm/2, -50, 50])
%axis tight
xlabel('Frecuencia [Hz]');
title('Espectro Se?al Base');
grid on;

```

Figura 8. Código en Matlab® para el análisis básico de frecuencia y fase de la señal de audio del sonido *C-flauta.wav*.

La gráfica de la Figura 9 muestra la señal en el tiempo y el espectro en frecuencia en hertz y la fase en radianes para la flauta en Do central.

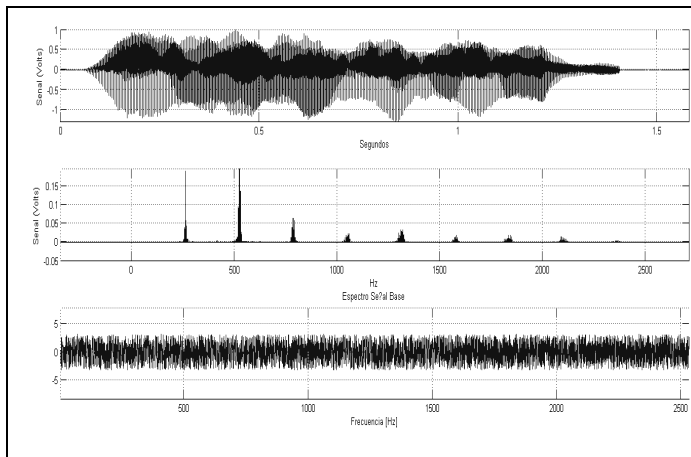


Figura 9. Forma de onda y espectro en amplitud y fase para la flauta en Do central.

La Figura 10 considera un detalle de aproximación sobre las curvas presentadas en la Figura 9 para facilitar la extracción de la información de amplitud A_n , frecuencia f_n y fase Φ_n .

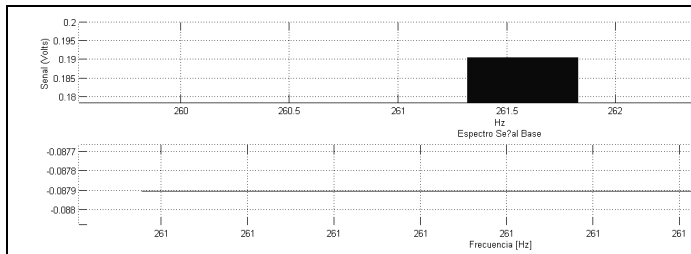


Figura 10. Ampliación de un tramo del espectro en magnitud y fase de la flauta en Do central.

La tabla 1 resume los valores leídos por inspección de la figura 9. Nótese que las frecuencias f_n no son múltiplos enteros de la frecuencia fundamental $f_1=261\text{Hz}$, pero se acercan; por lo que se denominan parciales.

Parcial n	Frecuencia f_n (Hz)	Amplitud A_n (Volts)	Fase Φ_n (Rads)
1	261.00	0.19	0.09
2	523.00	0.22	0.17
3	785.00	0.07	0.18
4	1050.00	0.02	2.60
5	1310.00	0.03	-0.90
6	1575.00	0.01	-1.74

Tabla 1. Valores leídos por inspección de la FFT del sonido de la flauta en Do central para los primeros 6 parciales.

Procediendo con la reconstrucción, se toman las amplitudes A_n , las frecuencias ω_n ($\omega_n=2\pi f_n$) y las fases Φ_n de la flauta desde la tabla 1 para posteriormente reconstruirle mediante la ecuación 2.

La Figura 11 entrega el código en Matlab® que permite su reconstrucción.

```
%Reconstruye
clear all
%Frecuencias
F1=261; F2=523
F3=785; F4=1050
F5=1310; F6=1575
%Amplitudes
A1=0.19; A2=0.22
A3=0.065; A4=0.02
A5=0.03; A6=0.01
%Fases
phi1=0.09; phi2=0.1685
phi3=0.184; phi4=2.6
phi5=-0.9042; phi6=-1.738
%Frecuencia de muestreo
fm=8000
T=1/F1
% Intervalo de tiempo
t=[0:T*F1/fm:T*F1]
%Serie de Fourier con 6 parciales
flauta=A1*cos(2*pi*F1*t+phi1)+A2*cos(2*pi*F2*t+phi2)+...
A3*cos(2*pi*F3*t+phi3)+A4*cos(2*pi*F4*t+phi4)+...
A5*cos(2*pi*F5*t+phi5)+A6*cos(2*pi*F6*t+phi6);
```

Figura 11. Código para reconstruir el sonido de la flauta.

Esta reconstrucción, hasta ahora, no ha considerado el hecho de que la señal real está sujeta a una envolvente; considerando esta situación se puede aplicar a la flauta la denominada ADSR (Attack, Decay, Sustain, Release).

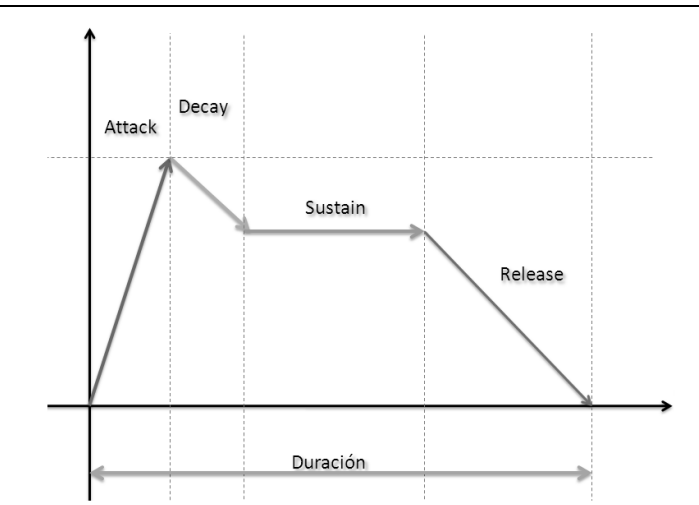


Figura 12. Envolvente ADSR clásica.

La envolvente puede depender del músico que interprete el instrumento de viento o de la naturaleza del instrumento, por lo cual la consideración de duración del cada ciclo presentada como ADSR en la reconstrucción es una aproximación para la flauta hecha por los autores (véase la figura 12).

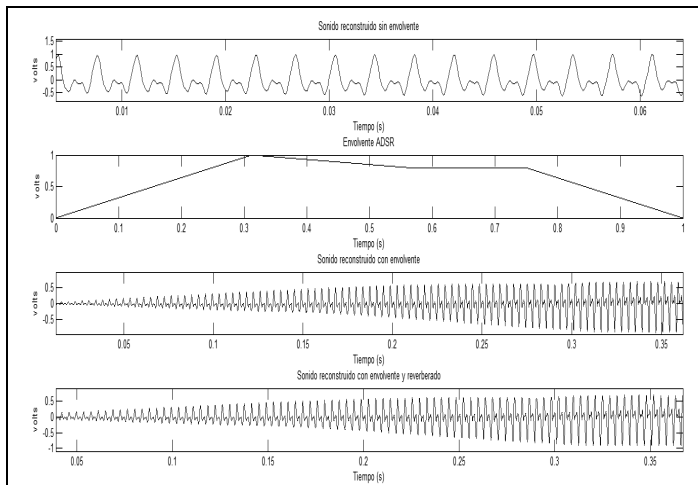


Figura 13. Señal reconstruida, Envolvente ADSR, señal con envolvente y reverberada.

Los instrumentos de cuerda como el piano no usan la envolvente ADSR, pues su principio físico es diferente a los instrumentos de viento. **La envolvente para una nota de piano carece de la etapa "Sustain"**, pasando directamente del "Decay" hacia el "Release". Sin embargo, los músicos experimentados manejan a su antojo la duración del "Release", así como la amplitud del "Attack", generando una amplia gama de sonidos. En la Figura 14 se muestra el código en Matlab® para normalizar la señal de la flauta reconstruida a 1 volt, crear la envolvente ADSR, producir una reverberación y guardar en un archivo .wav de salida.

```
%Normalizacion
flauta=flauta/max(flauta);
%Contruccion de la envolvente
%ADSR (Attack Decay Sustain Release).
t0=[0:2500]
ataque=t0*1/2500 %#ok<*NOPTS>
t1=[0:2000]
decaer=(-0.2/2000)*(t1-2000)+0.8
t3=ones(1,1498)
sostenido=t3*0.8
t4=[0:2000]
liberar=(-0.8/2000)*(t4-2000)
envolvente=[ataque decaer sostenido liberar]
falutaenv=flauta.*envolvente

% Efecto reverberante
flautarev=mi_reverb(falutaenv,0.5,0.2,fm)
% Sonido solo envolvente
wavplay(falutaenv,fm)
%Sonido reverberado mas envolvente
wavplay(flautarev,fm)
wavwrite(flauta,'C-mi_flauta.wav')
```

Figura 14. Composición por tramos de la envolvente ADSR en Matlab® para la flauta.

La reverberación brinda una sensación más realista, ya que esta introduce la superposición de una onda de sonido similar

pero retrasada en el tiempo debido al choque de esta en un espacio cerrado alcanzado, finalmente, al oyente. En la Figura 15 se entrega la función básica en Matlab® que permite reproducir un efecto reverberante con un tiempo de retraso t .

```
function salida=mi_reverb(entrada,gain,t,fm)
% Entrada: Sonido de entrada
% gain: Ganancia del reflejo
% t: Tiempo de retraso
% fm Frecuencia de muestreo
N=round(t*fm);
salida=entrada;
for n=N+1:length(entrada);
    salida(n)=entrada(n)+gain*entrada(n-N);
end
```

Figura 15. Función elemental en Matlab® para producir un efecto reverberante básico.

III. METODOLOGÍA

La presente investigación posee una metodología de tipo descriptiva para las técnicas básicas de síntesis de instrumentos musicales. Este trabajo ha sido realizado con estudiantes del curso de tratamiento digital de señales, quienes en el proceso han realizado:

- Consultas de las técnicas básicas de síntesis de la música dentro de la literatura científica.
- Implementación de algoritmos en Matlab® 7.0.
- Generación de los sonidos de la nota Do central y comparación cualitativa con un sonido patrón.

IV. RESULTADOS

La síntesis de la música por el método aditivo resulta ser de mediana complejidad a la hora de su implementación cuando se trata de instrumentos de viento. Ésta permite reproducir sonidos de la flauta y el saxofón de manera simple. **Para el caso de instrumentos como el piano el problema se hace más complejo ya que la frecuencia de la nota varía ligeramente en el tiempo** y su envolvente se ajusta más al modelo **ADR**. Durante 2 años un grupo de investigadores de la Universidad Tecnológica de Pereira han usado esta estrategia pedagógica para motivar los estudiantes en el estudio del tratamiento de las señales mediante la síntesis de la música.

V. CONCLUSIONES

El estudio de la síntesis de los sonidos de los instrumentos musicales es un tema motivador que interesa a los estudiantes de ciencias básicas, ingenierías y música ya que entrega una aproximación al tema desde diferentes perspectivas. El tema es altamente complejo y puede ser atacado desde la física del tubo para los instrumentos de viento o de la cuerda para otros. Se considera que es necesario continuar explorando este tema ya que en la literatura se encuentran trabajos de alto grado de complejidad con los cuales se puede motivar a los estudiantes a profundizar más en este aspecto para así trabajar,

posteriormente, la síntesis en sistemas embebidos para la música o la voz.

REFERENCIAS

- [1] Jesús Mariano Merino de la Fuente, “Las vibraciones de la música”, Editorial Club Universitario; Primera Edición (Julio 1, 2006).
- [2] Jaramillo Ana María, “Acústica la ciencia del sonido”, editorial ITM; (Enero 1, 2000)
- [3] Charles K. Alexander, y Matthew N. O. Sadiku, Fundamentos de Circuitos Eléctricos. McGraw-Hill Interamericana. México 2002.
- [4] J.A Cortes, F.A. Medina, J.A Chavés, “Del Análisis de Fourier A Las Wavelets Análisis De Fourier”, [En Línea] Disponible en: <http://revistas.utp.edu.co/index.php/revistaciencia/article/download/5563/3125> (consultado el 11 de junio de 2012)
- [5] J.A Cortes, J.A. Mendoza, J.A Muriel, “Alternativa al Análisis en Frecuencia de la FFT Mediante El Algoritmo Goertzel”, [En Línea] Disponible en: http://redalyc.uaemex.mx/redalyc/html/849/84917316040/84917316040_1.html 3125 (consultado el 11 de junio de 2012)
- [6] The Sonic Spot, “Types of Synthesis”, [En Línea] Disponible en: <http://www.sonicspot.com/guide/synthesistypes.html> 3125 (Consultado el 11 de junio de 2012)