

Apuntes sobre sonido digital

Por Antonio Sacco

Nota: estos apuntes fueron escritos por el profesor Antonio Sacco como complemento de sus clases dictadas sobre el tema. Para ampliar, referirse a las notas tomadas en clase y a la bibliografía de la materia. Año 2003-2004

Introducción	1
Algunos conceptos introductorios sobre el sonido	1
Cómo se almacena el sonido en una computadora	3
Principales diferencias entre el sonido digital y el almacenado analógicamente	5
Qué son los archivos WAV, MP3 y MIDI	6
Componentes de hardware necesarios para manejar sonido digital	7
Procesamiento digital del sonido	7

Introducción

Las computadoras actuales permiten, en general, manejar sonido digital. O sea que pueden almacenar, procesar y reproducir sonidos de diversos tipos.

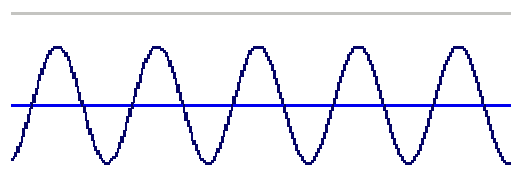
Las ventajas del procesamiento digital del sonido frente a otras técnicas más antiguas son muchas. En este pequeño apunte nos concentraremos en dos de las más importantes, que son a su vez comunes al procesamiento digital de todo tipo de información (imágenes, texto, video). Estas son: que el sonido almacenado digitalmente puede guardarse, copiarse y reproducirse infinitas veces sin perder calidad progresivamente; y que el sonido puede procesarse digitalmente de manera mucho más flexible y potente que mediante procesos analógicos.

Algunos conceptos introductorios sobre el sonido

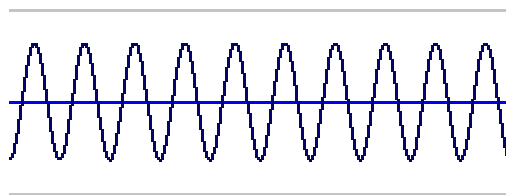
Para el propósito de este apunte, definiremos el *sonido* como la interpretación que hace nuestro cerebro de las variaciones de presión que genera un objeto vibrante en determinado medio (habitualmente, el aire). Para que esta vibración sea audible para un ser humano, el objeto debe oscilar aproximadamente entre 20 y 20.000 veces por segundo (ya que las frecuencias de ondas que somos capaces de escuchar van de 20 Hz -Hertz- a 20 KHz -kilo Hertz-).

Para describir un sonido “musical” se utilizan tres términos: altura, timbre e intensidad.

La *altura* está directamente relacionada con la frecuencia de oscilación, aunque algunos sonidos, como los percusivos, no tienen una altura definida. Lo que hace que un sonido posea o no una altura clara es la *periodicidad* de la señal. Es decir, que para que lleguemos a percibir su altura es necesario que la frecuencia de oscilación (que se mide en Hz -cantidad de ciclos de la señal por segundo-) no varíe dentro de un determinado lapso mínimo de tiempo. Mientras más alta sea la frecuencia de la señal, más agudo será el sonido, y viceversa. Por ejemplo, una señal sinusoidal de 440 Hz es un “La” musical. Si variamos la frecuencia al doble (880 Hz), obtendremos el La de la octava siguiente.

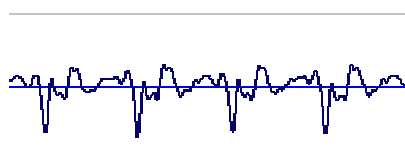


(sonido más grave)

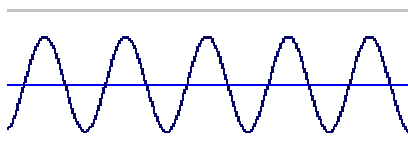


(sonido más agudo)

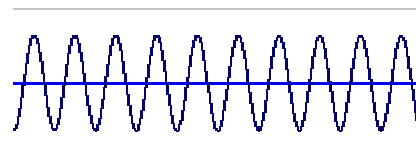
El *timbre* es lo que nos ayuda a caracterizar y distinguir diferentes tipos de instrumentos, o a reconocer a las personas por su voz, por ejemplo. Dos señales pueden tener la misma frecuencia (altura) pero diferente forma (timbre). Esto se debe a que las ondas de los sonidos naturales son complejas porque vibran con varias frecuencias simultáneas. La frecuencia de vibración más grave es la que determina normalmente el período y la altura, y se denomina *frecuencia base*. Las restantes frecuencias que suelen ser múltiplos de la frecuencia base se denominan armónicos, y son los que otorgan las “señas de identidad” a cada instrumento. O sea que cuando una señal compleja es periódica, ésta puede descomponerse como la suma de varias señales sinusoidales *simples*.



(sonido complejo)

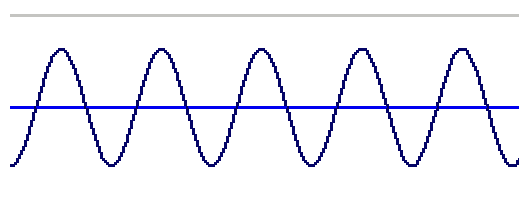


(componente simple)

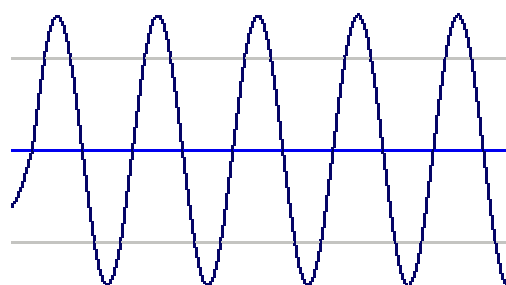


(componente simple)

La *intensidad* depende de la diferencia entre las presiones máxima y mínima que la onda puede alcanzar. Podemos asociarla al volumen del sonido, pero la percepción de la intensidad es un fenómeno auditivo muy complejo, por lo cual indicaremos (teniendo en cuenta que estamos simplificando) que mientras mayor sea la amplitud de la onda, mayor será el volumen y viceversa. Para indicar el nivel de intensidad se utiliza una escala logarítmica: los dB (decibelios).



(volumen más bajo)



(volumen más alto)

Cómo se almacena el sonido en una computadora

Lo primero que debemos estudiar para comprender qué es el sonido digital es el proceso que permite “introducirlo” dentro de una computadora.

Como sabemos, cualquier tipo de información que sea manejado por una computadora es almacenado en ésta mediante ceros y unos (información digital). Esos ceros y unos pueden representar texto, imágenes, sonido o cualquier otro tipo de elemento factible de ser procesado digitalmente, como por ejemplo un modelo tridimensional de una casa.

El proceso de *introducir* en una computadora cualquier elemento externo a ella se denomina *digitalización*. Un scanner, por ejemplo, es un dispositivo que permite digitalizar una imagen.

Para digitalizar el sonido se pueden utilizar diversos dispositivos, pero la solución más económica, en el caso de una PC, generalmente consiste en utilizar una placa de sonido y, eventualmente, un micrófono conectado a ella.

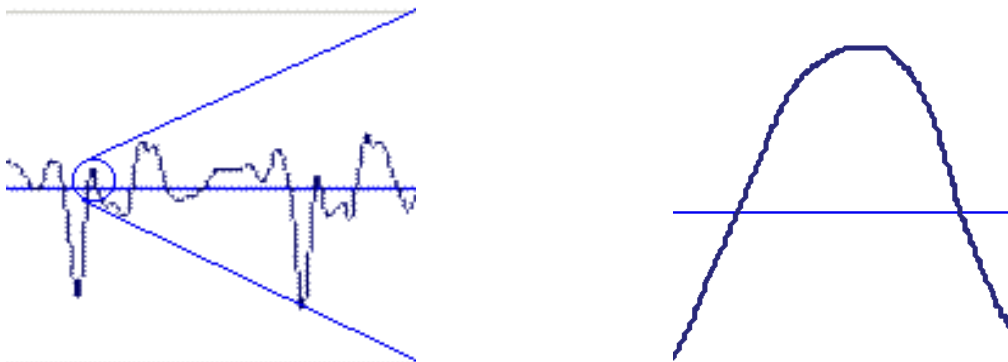
El micrófono convierte la variación de la presión de aire ejercida sobre su membrana en una señal eléctrica analógica (continua) que varía en el tiempo. Esta señal es introducida en la computadora, por ejemplo, mediante el cable que conecta el micrófono a la entrada correspondiente de la placa de sonido.

El dispositivo digitalizador (en nuestro caso la placa de sonido de la PC) *muestra* la señal analógica que recibe, una gran cantidad de veces por segundo.

La palabra *muestreo* es el equivalente del término inglés *sampling*, y se utiliza para indicar la acción de tomar muestras a intervalos de tiempo regulares. O sea que la digitalización en este caso consiste en “escuchar” el sonido de un instante muy breve (muestra), muchísimas veces por segundo, y almacenar una secuencia de dígitos binarios para cada muestra.

Luego, el proceso contrario, de reproducción del sonido almacenado, consiste en leer estas secuencias de ceros y unos y reproducir el sonido de cada muestra, tan rápido como cuando se las tomó.

Gráficamente, podemos representar este proceso de la siguiente manera:



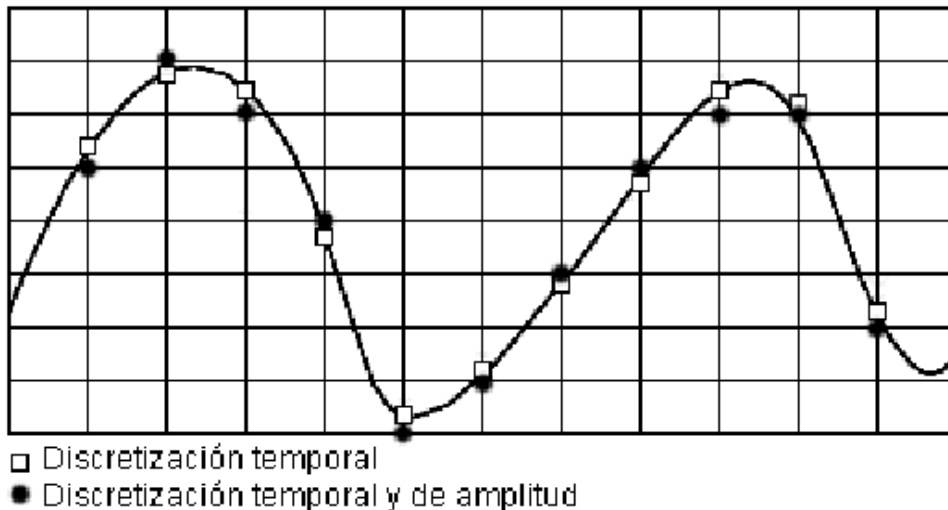
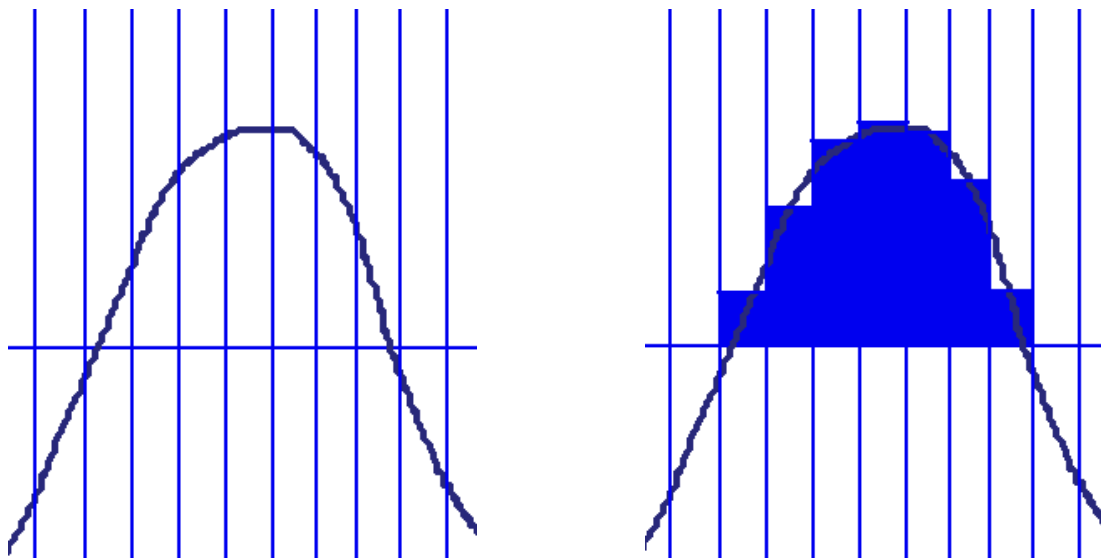


Figura 2.1. Discretización de una señal continua.

La *frecuencia* de muestreo es un número que indica la cantidad de muestras que se toman en determinado intervalo de tiempo; por ejemplo, en un segundo.

La *resolución* o *profundidad* del sonido es un número que indica cuántos bits (dígitos binarios, ceros y unos) se utilizan para representar cada muestra.

Tanto la frecuencia como la resolución están directamente relacionadas con la calidad del sonido digital almacenado. Mientras mayores sean estos indicadores, más parecida será la calidad del sonido digitalizado con respecto al real.

El estándar definido cuando se crearon los discos compactos de audio especifica que el sonido digital almacenado en ellos debe poseer una frecuencia de 44.1 KHz (KiloHertz) y 16 bits estéreo. Esto significa que se deben tomar unas 44100 muestras por segundo, cada una se representará con 16 bits, y en dos canales independientes (sonido estéreo).

Por lo tanto, si queremos saber cuántos ceros y unos se necesitan para almacenar 3 minutos de sonido digital con calidad de CD, debemos realizar el siguiente cálculo:

3 minutos x 60 segundos = 180 segundos.

180 segundos x 44100 muestras por seg. = 7.938.000 muestras.

7.938.000 muestras x 16 bits = 127.008.000 bits.

127.008.000 bits x 2 canales = 254.016.000 bits.

Entonces, para almacenar una canción de 3 minutos con calidad de CD necesitaremos 254.016.000 ceros y unos o, lo que es igual, 31.752.000 bytes (unos 30 Mb).

La magnitud de este resultado nos indica la gran cantidad de espacio de almacenamiento requerido para manejar sonido digital, y esta es una de las razones por las que no hace muchos años que esto es posible con las PCs de bajo costo.

Al contrario que el sonido, el texto no requiere una gran cantidad de espacio para ser almacenado en una computadora, y ello se hace evidente si pensamos que en el mismo disco compacto que es capaz de albergar cientos de miles de páginas de texto, apenas podemos guardar poco más de una hora de sonido digital de buena calidad. Más adelante observaremos que existen formas de reducir considerablemente el espacio necesario para almacenar sonido de buena calidad.

Principales diferencias entre el sonido digital y el almacenado analógicamente

Hay diversas formas de almacenar y reproducir el sonido sin utilizar tecnologías digitales. Por ejemplo, los cassettes tradicionales utilizan una cinta magnetizada, en la cual el cabezal de la grabadora “imprime” marcas magnéticas que representan las variaciones continuas de la señal analógica, las cuales luego son leídas y convertidas en electricidad nuevamente por el cabezal de reproducción, para ser enviadas a los parlantes y producir las vibraciones del sonido grabado.

En el caso de los discos de vinilo, éstos poseen impreso un surco en forma de espiral, que al ser recorrido por la púa la hace vibrar y estas vibraciones son amplificadas para que escuchemos el sonido allí almacenado.

Estas dos formas de grabación de sonido (y otras que no comentaremos aquí) se denominan analógicas.

Una forma sencilla de comprender el término *analógico* en este contexto (aunque no es la más prolija desde un punto de vista estrictamente teórico) es por contraposición con el término *digital*. Llamaremos representaciones digitales a todas aquellas que refieran algún estado o proceso del mundo real mediante ceros y unos (dígitos binarios). Luego, la mayoría de las demás representaciones del mundo real que no utilizan este sistema, en general serán analógicas. La información analógica es continua (se pueden medir infinitos valores), mientras que la información digital es discreta (mientras más ceros y unos utilicemos, más valores se pueden medir, pero siempre es una cantidad limitada).

En la introducción de este apunte comentamos que una de las ventajas más importantes del almacenamiento digital del sonido es que éste no pierde calidad al copiarlo, reproducirlo o simplemente conservarlo durante largo tiempo. En cambio, las cintas de cassettes y los discos de vinilo se deterioran progresivamente con el uso, y este deterioro físico tiene como consecuencia una pérdida también progresiva en la fidelidad del sonido almacenado.

¿Por qué no sucede esto con el sonido digital? Porque una determinada cadena de unos y ceros no cambiará luego de copiarla o reproducirla varias veces. Cabe aquí la aclaración de que el medio de almacenamiento de la información digital sí puede sufrir un pequeño desgaste (por ejemplo el deterioro del disco compacto, disco rígido, disquette, etc.), pero este es insignificante en comparación al de los medios analógicos y, además, no genera una modificación progresiva de la información almacenada, sino fallas puntuales que pueden ser prevenidas convenientemente.

Otra gran ventaja del almacenamiento digital del sonido es que admite una amplia variedad de procesos para alterarlo voluntariamente. Existen programas de computación y dispositivos específicos que permiten mejorar o simplemente modificar de diversas formas el

sonido digital almacenado. Algunos de estos procesos se pueden lograr también por medios analógicos, pero los digitales suelen ser más variados y potentes.

Debemos tener en cuenta que una grabación analógica no tiene por qué sonar peor que una grabación digital, ni la digital tendrá necesariamente mejor calidad que una analógica, pero en la gama de reproductores de audio domésticos en general obtendremos mejores resultados con las grabaciones digitales y, como ya comentamos, la degradación de los medios analógicos es mucho más importante.

Qué son los archivos WAV, MP3 y MIDI

Cuando digitalizamos sonido para almacenarlo en el disco de una computadora debemos elegir el formato de archivo que utilizaremos. El formato estándar de Windows para almacenamiento de sonido se llama WAV.

Los archivos WAV almacenan todos los bits obtenidos de la digitalización, por lo cual si queremos calcular el espacio que ocupará un archivo de este formato debemos considerar la duración del sonido, la frecuencia de muestreo, la resolución utilizada, y luego realizar un cálculo como el que se ejemplificó antes en este mismo apunte. Como consecuencia, estos archivos ocupan mucho lugar en el disco pero conservan toda la calidad obtenida en la digitalización.

Debido a la necesidad de manejar archivos de sonido más pequeños (en cuanto a su tamaño en bytes) se diseñaron otros formatos de almacenamiento, de los cuales uno de los más populares es el MP3. Este formato aprovecha el hecho de que ciertos sonidos no son habitualmente distinguidos por el oído humano, y a esto le suma la utilización de complejas técnicas de compresión para lograr archivos varias veces más pequeños que los correspondientes WAV, en los cuales la diferencia de calidad es prácticamente imperceptible.

Para obtener un archivo MP3 a partir de uno WAV hay que utilizar el software adecuado e indicar varios parámetros que serán tenidos en cuenta en la conversión, de los cuales dependerán la calidad y el tamaño del archivo obtenido. Existen otros formatos de sonido similares que no explicaremos aquí, por escapar a los alcances de este apunte.

Tanto los archivos WAV como los MP3 almacenan el sonido sin importar de qué se trate. Si se los utiliza para grabar una canción en la cual participan una guitarra y un piano, si bien el hombre podrá distinguir ambos instrumentos, para la computadora es un solo sonido formado por la suma de ambos (no consideramos aquí la posibilidad, real de hecho, de grabar cada instrumento en un *canal* separado).

En cambio, existe un formato denominado MIDI que almacena lo que podríamos llamar “la interpretación” de los instrumentos, en vez del sonido real generado por ellos. Simplificando bastante el proceso, imaginemos que lo que se almacena en un archivo MIDI son las partituras interpretadas por cada instrumento, con las anotaciones necesarias para que la máquina sepa qué instrumento se debe utilizar en cada caso y varios detalles más. Estas “partituras” indican la nota a interpretar, su duración, etc., y de ello se desprenden ventajas y desventajas con respecto a los formatos WAV y MP3.

Una ventaja de los archivos MIDI es que son mucho más pequeños que los WAV y los MP3, ya que la cantidad de información necesaria para indicar qué notas deben ser interpretadas es menor que la necesaria para las muestras de los otros. Otra ventaja muy importante es que con este formato se pueden realizar modificaciones sustanciales en la interpretación al nivel de, por ejemplo, cambiar una nota o reemplazar un instrumento por otro.

La mayor desventaja del formato MIDI con respecto a los otros dos es que el primero nunca logra reproducir de manera totalmente fiel la interpretación original (si es que esta fue interpretada alguna vez por un músico con su instrumento, ya que puede haber sido compuesta

directamente sobre una computadora). Esto se debe a que cada vez que se reproduce un archivo MIDI, la computadora “interpreta” la melodía de acuerdo a su capacidad. Y esto hace que el mismo archivo reproducido por diferentes equipos pueda sonar diferente.

Componentes de hardware necesarios para manejar sonido digital

No haremos aquí un análisis de los distintos componentes que toman parte en el procesamiento de sonido digital; tan sólo enumeraremos los más importantes.

En primer lugar debemos dejar en claro que nos referiremos a un procesamiento de sonido relativamente simple con una computadora personal, no a lo que se necesita para un trabajo profesional, que en general requerirá de equipos especiales mucho más costosos.

Necesitaremos una PC con una tarjeta de sonido (actualmente todas las PC ya traen tarjeta de sonido). La tarjeta de sonido es el componente encargado de convertir la información digital en una señal eléctrica que luego es enviada a los parlantes para ser convertida finalmente en vibraciones (el sonido que escuchamos). Eventualmente, la tarjeta de sonido también es la encargada de realizar el proceso contrario, o sea la digitalización del sonido explicada anteriormente.

Necesitaremos parlantes o auriculares y, en algunas ocasiones, podemos usar un amplificador entre la placa de sonido y los parlantes, para lograr una mayor potencia.

Para digitalizar nuestra voz o el sonido del ambiente podemos utilizar un micrófono. Muchas placas de sonido admiten conectarles, como entrada, la salida de algún otro equipo de audio y, en general, la misma lectora de CDs puede utilizarse para leer el sonido de un disco compacto.

La capacidad de manejar el formato MIDI no es intrínseca de cualquier placa de sonido, aunque la mayoría de ellas lo permiten en la actualidad. Si queremos ejecutar una pieza en un instrumento y que la máquina vaya *grabando* nuestra interpretación en formato MIDI, necesitaremos conectar la placa de sonido (con entrada para MIDI) de nuestra computadora a un instrumento con *interfaz* MIDI. Por ejemplo, si nuestro teclado tiene una ficha para “salida MIDI”, al conectarlo a la computadora y tocar una pieza, la PC registrará las notas que tocamos, su duración, etc.

Procesamiento digital del sonido

Una vez que tenemos guardado en el disco de una PC un archivo de sonido, podemos “retocar” y aplicar diferentes efectos, generalmente llamados *filtros*. Es posible realizar desde operaciones sencillas, como quitar una porción o repetir otra, modificar el volumen, etc., hasta otras más complejas, como destacar ciertos matices, aplacar otros, etc.

Existe una amplia gama de software para procesamiento de archivos de sonido, desde la sencillísima “Grabadora de sonido” de Windows, pasando por el CoolEdit o el GoldWave, hasta programas más complejos como el SoundForge.

El programa GoldWave tiene la ventaja de ser shareware (se puede descargar gratuitamente de la dirección de Internet www.goldwave.com) y permite realizar una gran cantidad de operaciones con sonido digital (suficientes para obtener excelentes resultados en un ambiente no profesional). Además, posee un completo sistema de ayuda.

Otro buen programa es “Audacity”, un editor de audio gratuito. Permite grabar y reproducir sonidos, importar y exportar archivos WAV, AIFF, y MP3, entre otros, editar usando Cortar, Copiar y Pegar (con ilimitados Deshacer), mezclar pistas, o aplicar efectos a las

grabaciones. Audacity es desarrollado por un grupo de voluntarios bajo el modelo de código fuente abierto (open-source), por lo que el código fuente del programa está disponible bajo la licencia GNU General Public License, la cual esencialmente permite a cualquiera modificar el código fuente mientras publique los cambios. La página web para descargarlo, consultar manuales y actualizarlo es: <http://audacity.sourceforge.net>