

Introducción al sonido digital

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. Escrito en 2003, revisado en 2015.

<i>Introducción</i>	<u>1</u>
<i>Algunos conceptos introductorios sobre el sonido</i>	<u>1</u>
<i>Cómo se almacena el sonido en una computadora</i>	<u>3</u>
<i>Principales diferencias entre el sonido digital y el almacenado analógicamente</i>	<u>5</u>
<i>Archivos WAV, MP3 y MIDI</i>	<u>6</u>
<i>Componentes de hardware necesarios para manejar sonido digital</i>	<u>7</u>
<i>Procesamiento digital del sonido</i>	<u>8</u>
<i>Analogías entre imágenes y sonido digital</i>	<u>8</u>
<i>Fuentes consultadas</i>	<u>11</u>

Introducción

Las computadoras permiten almacenar, procesar y reproducir sonido digital.

Las ventajas del procesamiento digital del sonido frente a técnicas analógicas 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 copiarse y reproducirse sin perder calidad progresivamente; y que el sonido puede procesarse digitalmente de manera mucho más flexible y potente que mediante procesos analógicos.

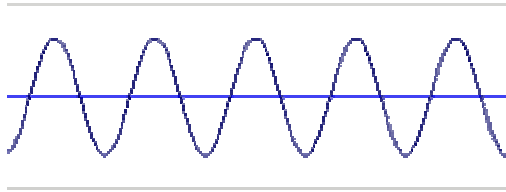
Algunos conceptos introductorios sobre el sonido

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-*).

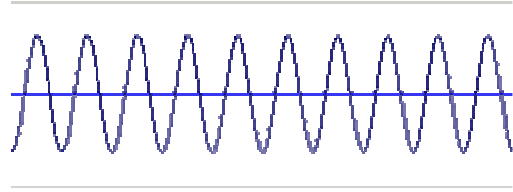
Para describir un sonido "musical" se utilizan tres términos: altura, timbre e intensidad.

La *altura* está directamente relacionada con la frecuencia de oscilación, aunque algunos sonidos, como los percusivos, no siempre 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.

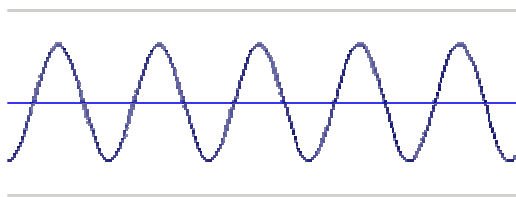


(sonido más grave)

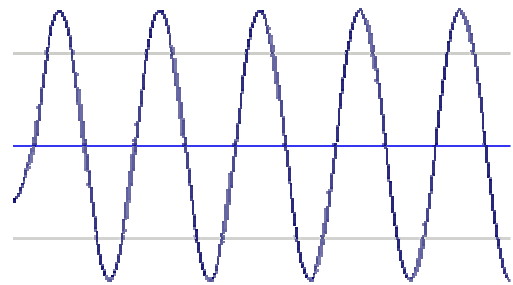


(sonido más agudo)

La *intensidad* depende de la diferencia entre las presiones máxima y mínima que la onda puede alcanzar. Podemos asociarla al "volumen" del sonido, aunque 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).

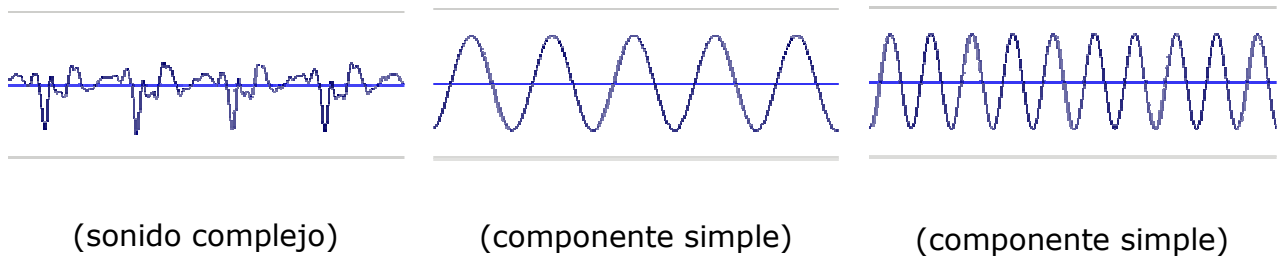


(volumen más bajo)



(volumen más alto)

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*.



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 representado en ésta mediante ceros y unos (información digital). Esos ceros y unos pueden representar texto, imágenes, sonido o cualquier otro tipo de información.

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, como por ejemplo una PC con placa de sonido y, eventualmente, un micrófono conectado a ella.

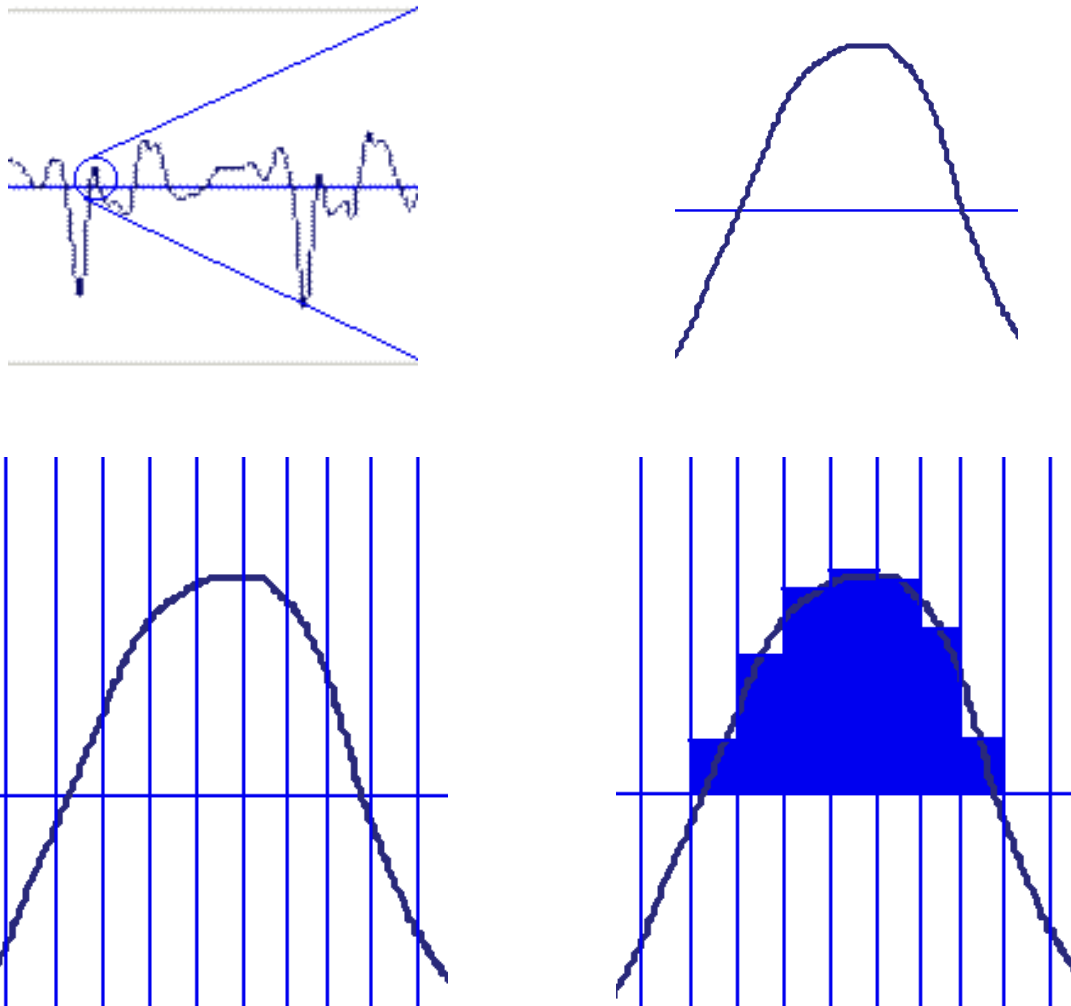
El micrófono convierte la variación de la presión de aire ejercida sobre su membrana en una señal eléctrica analógica (continua) que varía en el tiempo. Esta señal es introducida en la computadora, por ejemplo, mediante el cable que conecta el micrófono a la entrada correspondiente de la placa de sonido.

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:



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 *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 cantidad de bits por muestra 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:

$$3 \text{ minutos} \times 60 \text{ segundos} = 180 \text{ segundos}$$

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

254.016.000 bits / 8 bits por byte = 31.752.000 bytes

Entonces, para almacenar una canción de 3 minutos con calidad de CD necesitaremos 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.

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.

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 esa vibración se convierte en electricidad, que luego es amplificada y convertida nuevamente en sonido.

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 utilizemos, más valores se pueden medir, pero siempre es una cantidad limitada).

En la primera parte de estas notas 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.

Pero, ¿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 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.

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 cantidad de bits por muestra, 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 información 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 percibidos 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, como 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 básicamente se utiliza para almacenar música instrumental, pero no admite la voz humana. A su vez, el MIDI 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 creada 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 a 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 "retocarlo" 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 completos como el "Audacity". Este último es un editor y grabador de audio multipista libre. Permite grabar y reproducir sonidos, importar y exportar archivos WAV, AIFF, y MP3, entre otros, editar usando Cortar, Copiar y Pegar, mezclar pistas, aplicar efectos a las grabaciones, etc. La página web para descargarlo, consultar manuales y actualizarlo es: <http://audacity.sourceforge.net>

Analogías entre imágenes y sonido digital

La digitalización, almacenamiento y procesamiento de las imágenes y el sonido tienen varios puntos en común, y su estudio puede ayudar a clarificar algunos conceptos, por lo que mencionaremos brevemente algunos de los aspectos más parecidos de estos dos campos.

Digitalización

Digitalización de una imagen: para "introducir" una imagen dentro de la computadora tenemos que digitalizarla, esto es, convertirla en información digital (ceros y unos). Si tenemos la imagen en papel, una de las formas de digitalizarla es mediante un scanner. Este dispositivo "mira" el color de cada punto (muestra) del papel, le asigna un valor (código formado por ceros y unos) y envía esa información en cierto orden y con determinados datos de control a la computadora. Una vez que tenemos la imagen en la memoria de la computadora, podemos almacenarla en un disco, por ejemplo como un archivo con formato BMP.

Digitalización del sonido: para "introducir" el sonido dentro de la computadora, al igual que con las imágenes, tenemos que digitalizarlo. Si queremos grabar nuestra voz podemos usar un micrófono conectado a la placa de sonido de la computadora, a la cual también es posible conectar un equipo de música (reproductor de CDs, de cassettes, de discos, etc.). En cualquiera de estos casos, es esta placa la que "escucha" pequeñas fracciones (muestras) de sonido, les asigna un valor (código formado por ceros y unos) y envía esa información en cierto orden y con determinados datos de control a la computadora. Una vez que tenemos el sonido en la memoria de la computadora, podemos almacenarlo en un disco, por ejemplo como un archivo con formato WAV.

Resolución

Resolución de una imagen: podemos imaginar una imagen digital como una grilla formada por muchísimos pequeños puntos de colores. Mientras más pequeños sean esos puntos, mejor se verá la imagen y, a su vez, necesitaremos más información (ceros y unos) para representarla. La cantidad de puntos por unidad de espacio que conforman una imagen digital se denomina resolución, y en general se mide en puntos por pulgada (PPP, o DPI -por Dots Per Inch-). Si escaneamos una imagen con poca resolución (por ejemplo, 75 dpi), podremos verla más o menos bien en la pantalla con sus dimensiones originales, pero al intentar ampliarla comenzará a distorsionarse, ya que cada punto deberá ser "dibujado" con varios pixels (el punto más pequeño que muestra un monitor). Decimos en ese caso que la imagen se pixelará (en las curvas y diagonales se observará un efecto "serrucho"). Para poder ampliar la imagen sin distorsión debemos digitalizarla con una resolución mayor, como por ejemplo 200 dpi. La resolución también dependerá del medio en el que necesitaremos la imagen. Para imprimir una foto en una casa de fotografías con un tamaño de hasta 13x18 cm. es conveniente que tenga aproximadamente 300 dpi.

Resolución del sonido: al sonido digital, a su vez, lo podemos imaginar como una secuencia de muchísimos "instantes" de sonido (llamados muestras). Mientras más breves sean esas muestras, mejor se escuchará, y también necesitaremos más información (ceros y unos) para representarlo. La cantidad de muestras por unidad de tiempo que conforman un sonido digital se mide en muestras por segundo (Hz -por Hertz-). Si digitalizamos el sonido con pocas muestras por segundo (por ejemplo, 8000 Hz), podremos entender más o menos bien las voces, pero para obtener una mejor calidad que permita disfrutar los matices de muchos instrumentos deberemos usar una resolución mayor. Por ejemplo, el sonido de los CDs de audio tiene una resolución de 44100 Hz.

Profundidad

Profundidad del color de una imagen: dijimos que una imagen digital está formada por una larga serie de puntos, y el color de éstos está representado por un código formado por ceros y unos (bits). Ahora supongamos que para cada punto podemos usar solo un bit. Las posibilidades son dos (un cero o un uno) y, por lo tanto, cada punto podrá tener únicamente uno de dos colores (blanco o negro, azul o blanco, amarillo o rojo, etc.) Si en lugar de un bit utilizamos dos para cada punto, tendremos cuatro posibles colores (que serán representados con 00, 01, 10 y 11), y así sucesivamente. Queda claro que, mientras más bits por punto utilizemos, más colores podrá tener la imagen y, consecuentemente, mejor se verá. A la cantidad de bits por punto se la denomina profundidad de color, y algunos valores típicos son: 8 bits (que permite 256 posibilidades diferentes), 16 bits (para tener hasta 65536 colores), 24 bits (con más de 16 millones de tonalidades).

Profundidad del sonido: en el caso del sonido digital, cada muestra también debe ser codificada con una serie de ceros y unos que representan el "sonido" captado en el momento en que se tomó la muestra. Si consideramos que cada serie de ceros y unos representa la magnitud de la onda de sonido en un instante, podemos deducir que mientras más bits por muestra utilizemos, más posibilidades diferentes de sonido tendremos y, entonces, mejor será el resultado al escucharlo. El sonido con "calidad de CD", por ejemplo, es de 16 bits. O sea que cada una de las 44100 muestras de sonido que se toman por segundo es codificada con 16 ceros y unos.

Formatos

Formatos de archivos de imágenes: el formato BMP almacena absolutamente toda la información digitalizada; esto es, todos los ceros y unos que representan el color de cada uno de los puntos de la imagen. O sea que, incluso si guardamos en un archivo una imagen totalmente blanca, el tamaño (en bytes, Kb o Mb) será el mismo que si se trata de una compleja fotografía (siempre que las dimensiones, resolución y profundidad de color sean las mismas). Como el formato BMP ocupa mucho espacio en el disco (por lo que ya se comentó), existen formatos con pérdida de información, como por ejemplo el JPG y el GIF. JPG es ideal para fotografías, ya que el algoritmo de compresión que utiliza funciona muy bien con variaciones de color suaves, y admite millones de colores. GIF es un excelente formato para ilustraciones, ya que admite solo hasta 256 colores y funciona muy bien con cambios bruscos de color (lo que sucede típicamente cuando uno dibuja con la computadora, por ejemplo, usando un programa como el Paint).

Formatos de archivos de sonido: el formato WAV, análogamente al BMP de las imágenes, guarda todos los bits de cada muestra, independientemente de que todo sea "silencio", o haya sonidos más o menos complejos. Para contrarrestar el gran tamaño (en Mb) de los archivos WAV existen formatos de sonido con pérdida de información, como por ejemplo el MP3. El algoritmo de compresión del formato MP3 descarta información que no es perceptible por el oído humano para lograr que el mismo fragmento de audio pueda ocupar diez veces menos en el disco y se trata de un formato que podríamos considerar análogo al JPG en imágenes.

Formatos descriptivos

Gráficos vectoriales: existen formatos de imágenes digitales en los que no se almacena el color de cada punto, sino fórmulas que "explican cómo dibujar" la imagen. Uno de estos formatos, de los llamados vectoriales, se denomina WMF (por Windows Meta File). Las dos ventajas principales son: que su peso (en Kb) en general es mucho menor que el de los archivos BMP, JPG o GIF (llamados "rasterizados"), y que permite ampliar o achicar la imagen todo lo que se desee sin perder calidad (ya que aplicando una transformación matemática a las fórmulas descriptivas se obtienen "líneas" iguales pero más grandes, con las que se dibuja la imagen). Como desventaja, el formato WMF no sirve para fotografías (aunque puede incluirlas).

Archivos MIDI: este formato de archivos no almacena "sonido grabado", sino las indicaciones para que la computadora interprete una melodía. Así, podemos imaginarlo como algo similar a un conjunto de partituras, con los nombres de los instrumentos que hay que utilizar, las notas, tiempos y algunas aclaraciones acerca de la interpretación. Los archivos MIDI también tienen dos grandes ventajas con respecto a los de sonido grabado: que son muchísimo más pequeños, y que la melodía puede ser editada modificando las notas sobre un pentagrama con el software adecuado. Como desventaja, la calidad de la música dependerá de la "interpretación", que está asociada a la calidad de la placa de sonido, y el formato MIDI no sirve para grabar la voz.

Tamaños

Ejemplo para un archivo BMP: para saber cuánto ocupará una fotografía al digitalizarla y guardarla con formato BMP:

1. Multiplicamos la cantidad de píxeles de ancho por la cantidad de píxeles de alto de la fotografía
2. Multiplicamos la cantidad total de puntos (obtenida en el punto anterior) por la cantidad de bits utilizada para representar el color de cada punto. Esto nos da la cantidad total de bits de la fotografía.
3. Dividimos la cantidad total de bits de la foto (obtenida en el punto anterior) por 8, para obtener su tamaño en bytes (ya que cada byte está formado por 8 bits).

Por ejemplo, si tenemos una fotografía de 1000x480 píxeles y color de 24 bits:

- 1) $1000 \times 480 = 480.000$ píxeles
- 2) $480.000 \times 24 = 11.520.000$ bits en total
- 3) $11.520.000 / 8 = 1.440.000$ bytes en total (aprox. 1,4 Mb)

Ejemplo para un archivo WAV: para saber cuánto ocupará un sonido al digitalizarlo y guardarlo con formato WAV:

1. Obtenemos la duración total del sonido (en segundos)
2. Multiplicamos la duración (obtenida en el punto 1) por las muestras por segundo. Esto nos dará la cantidad total de muestras que conformarán el sonido al digitalizarlo
3. Multiplicamos la cantidad total de muestras del sonido (obtenida en el punto anterior) por la cantidad de bits. Esto nos da la cantidad total de bits del sonido.
4. Dividimos la cantidad total de bits del sonido (obtenida en el punto anterior) por 8, para obtener su tamaño en bytes (ya que cada byte está formado por 8 bits).

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:

- 1) $3 \times 60 = 180$ segundos
- 2) $180 \times 44100 = 7.938.000$ muestras en total
- 3) $7.938.000 \times 16 = 127.008.000$ bits
- 3.a) $127.008.000 \times 2 = 254.016.000$ bits en total (porque el sonido en el CD es estéreo)
- 4) $254.016.000 \text{ bits} / 8 = 31.752.000$ bytes en total (aprox. 30 Mb).

Fuentes consultadas

Jordà, Sergi. "Audio digital y MIDI". Ed. Anaya Multimedia, Madrid, 1997. (En formato PDF en: <http://www.tecn.upf.es/~sjorda/Anaya/AnayaPDFSAll.zip>, y como página web en: <http://www.ccapitalia.net/reso/articulos/audiodigital/introduccion.htm>)

Berenguer, J.M, M. Dalmases y S. Jordà, Sistemas acústicos y de tratamiento del sonido y del habla, Universitat Oberta de Catalunya, Barcelona 2001.

(http://cv.uoc.es/moduls/UW01_50001_00332/index.html,
http://cv.uoc.es/moduls/UW01_50001_00333/index.html,
http://cv.uoc.es/moduls/UW01_50001_00334/index.html,
http://cv.uoc.es/moduls/UW01_50001_00335/index.html)

Guías y tutoriales / Teoría sobre MIDI y Audio

(<http://www.electrofante.com/musicadigital>)

Introducción a los formatos digitales de sonido

(<http://www.electrofante.com/musicadigital/introduccion.html>)

Comparativa entre el MIDI y el audio digital

(http://www.electrofante.com/musicadigital/midi_vs_audio.html)