

---

---

# ***Sonido digital y formatos de compresión***

*José Manuel Huidobro, Ing. de Telecomunicación*

---

## **SONIDO DIGITAL Y FORMATOS DE COMPRESIÓN**

Como de todos es sabido, el sonido se produce por las vibraciones de las partículas de aire, vibración ocasionada por cualquier método físico, como es la voz, la percusión de un objeto o el paso de aire por un instrumento musical, que se propagan y que llegan a nuestros oídos, excitándolos con la presión que se ejerce sobre ellos. La velocidad de propagación de las ondas sonoras –acústicas– depende del medio en cuestión, y así en el aire es de 340 m/s mientras que en el acero llega a alcanzar los 5.000 m/s; en el vacío no se propagan, al contrario de lo que sucede con las ondas electromagnéticas.

La audición en los seres humanos ocurre siempre que una vibración tenga una frecuencia comprendida entre unos 15 y 20.000 hertzios, y su intensidad sea la suficiente para llegar al oído interno y excitarlo. Cuando las vibraciones superan este margen se habla de ultrasonidos y no son perceptibles para el ser humano, lo mismo que tampoco lo son las de frecuencia muy baja o infrasonidos.

Las características del sonido (onda sonora) se pueden medir y para ello se usan las unidades de hertzios

(Hz) que miden la frecuencia de un sonido, o sea cuantas veces vibra en un segundo (ciclos/segundo), y los decibelios (dB) que mide la intensidad (amplitud) de una onda. También, hay que considerar el timbre, una cualidad que diferencia las mismas frecuencias pero producidas por voces o instrumentos diferentes.

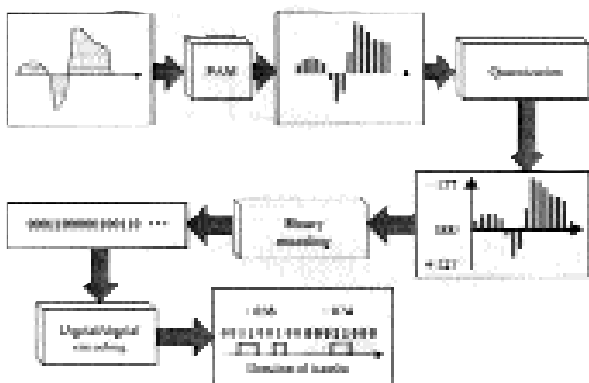
El sonido, en sus múltiples versiones, está presente en nuestra vida cotidiana y su tratamiento y almacenamiento en formato electrónico digital cada vez es más frecuente, utilizando herramientas informáticas (hardware y software) específicas, o comunes como son los PC. Se busca siempre un sonido de calidad y eso solo se puede conseguir con un formato digital, que garantiza de antemano los resultados que se van a obtener y permite su almacenamiento, distribución y reproducción sin pérdida de las características originales.

El oído y un micrófono incorporado a la tarjeta de sonido de un PC tienen un funcionamiento similar. Ambos transforman las vibraciones del aire en una señal eléctrica que puede ser comprendida y almacenada por el cerebro o la CPU, respectivamente. Esta señal, que es analógica, puede ser guardada, manipulada –por ejemplo, digitalizada– y reproducida por los medios electrónicos adecuados.

La palabra “digital” nos indica la presencia de procesos numéricos para concretar un hecho (imagen,



sonido, etc.). Los sistemas de audio digital tienen circuitos electrónicos para guardar el registro de la música; en pocas palabras, lo que hacen estos circuitos es grabar una larga cadena de números generados con un dispositivo llamado conversor analógico digital (A/D), que se encarga de muestrear la onda y asignarle a cada momento un valor numérico (cuantificación y codificación), un proceso que se conoce como **digitalización**; posteriormente ese valor numérico es decodificado por un conversor llamado digital-analógico (D/A) para convertirla otra vez en analógica y poder ser escuchada a través del altavoz reproductor.



Proceso de muestreo, cuantificación y codificación para convertir una señal analógica (sonido) en formato digital.

## CALIDAD DEL SONIDO DIGITAL

La calidad del sonido digital depende de la frecuencia de muestreo y de la resolución.

La **frecuencias de muestreo** (*sampling*), se refiere al número de mediciones que se hacen por segundo. Cuanto mayor sea el número de muestras mejor es la calidad del sonido; así, por ejemplo si la velocidad de muestreo es de una muestra por cada segundo las variaciones del sonido que se produzcan en el intermedio no serán registradas. Según estudios (Teorema de Nyquist), la frecuencia de muestreo debe ser el doble del sonido más alto que se pueda escuchar; por tanto, como el oído humano puede escuchar aproximadamente hasta los 20.000 hertzios, la frecuencia óptima de muestreo será de 40 kHz, pero se ha elegido 44,1 kHz (44.100 hertzios) para tener un cierto margen de seguridad, que es la frecuencia que se usa en los CD de música. Como los instrumentos o las voces humanas no pasan de a frecuencia de los 10 kHz., con una frecuencia de muestreo inferior, de 32 kHz sería mas

que suficiente, por lo que también se utiliza este valor. Si se bajase mucho más la frecuencia de muestreo el sonido se vuelve opaco o poco nítido pues se pierden las frecuencias agudas.

**La resolución** es el término que hace referencia a la exactitud de las medidas de frecuencia y se mide en número de bits. Si la resolución es de 8 bit tenemos 256 niveles posibles ( $2^8=256$ ). Si se amplía a 16 bit, el rango se extiende a 65.535 ( $2^{16}=65.536$ ). Como referencia se puede decir que un disco compacto se graba a 44,1 kHz y a una resolución de 16 bits (margen dinámico de 96 dB).

Por tanto, si se desea digitalizar 5 minutos de música, a un muestreo de 44,1 kHz y almacenando por cada muestra dos bytes (16 bits), se obtiene lo siguiente:

$$5 \text{ min.} \times 60 \text{ seg./min.} \times 44.100 \text{ muestras/seg.} \times 2 \text{ bytes/muestra} = 26.460.000 \text{ bytes}$$

El cálculo nos indica que para almacenar una canción de 5 minutos con calidad profesional se necesitarán 26 MB aproximadamente. Es aquí donde surge el problema, pues aunque la capacidad de los discos duros (HD) u otros medios de almacenamiento ha crecido mucho, tener varias canciones significaría ocupar gran parte del disco duro del ordenador. Para solucionar este problema se han desarrollado formatos de archivo que permiten realizar grabaciones de sonido con muy buena calidad usando un método de compresión que disminuye considerablemente el número de bytes ocupados; el problema es que estando así el sonido no puede ser editado para ser modificado.

## FORMATOS DE SONIDOS MÁS POPULARES

A continuación se da una breve descripción de los formatos mas usados para el sonido digitalizado (secuencia de ceros y unos):

### Formato MIDI

El formato MIDI proviene de Musical Instrument Digital Interface (Interface digital para instrumentos musicales). Es un protocolo de comunicación estándar utilizado para combinar datos entre sintetizadores, procesadores de efectos y otros dispositivos MIDI.

Este es el formato mas usado en la composición musical y tiene generalmente la extensión mid. El



archivo contiene información de secuenciado, es decir, acerca de cuando tocar que instrumento y de que forma. Dependiendo del hardware, el sonido puede ser excelente o bien muy por debajo de lo aceptable.

Los sonidos (timbres) de los diferentes instrumentos tienen un número de programa y van desde el 1 al 128; generalmente se asigna el 1 al piano, además cada programa tiene parámetros propios como por ejemplo con qué intensidad atacar un sonido.

Toda la información MIDI que puede procesar un teclado sale como "Midi Out" y entra por el "Midi In" del ordenador, donde es completamente reconocido, interpretado y convertido en números. Se puede editar y manipular gracias al software apropiado.

mido y grabado en distintas calidades y tamaños (desde 11.025 hasta 44.100 kHz).

Aunque los archivos Wav pueden tener un excelente sonido comparable al del CD (16 bytes y 44,1 kHz. estéreo) el tamaño necesario para esa calidad es demasiado grande (especialmente para los usuarios de Internet). Una canción convertida a Wav puede ocupar fácilmente entre 20 y 30 MB. Una opción que reduce el tamaño es grabar a 4 bits y con pocos kHz; el problema es la baja calidad del sonido, los ruidos, la estática, incluso cortes en el sonido, por esta razón casi siempre se usa solo para muestras de sonido. La ventaja más grande es la de su compatibilidad para convertirse en varios formatos por medio del software adecuado, por ejemplo pasar de Wav a MP3.



*Reproductor Microsoft de Windows, para formato multimedia.*

## Formato MP3

Este formato de compresión de audio fue creado por el Moving Picture Expert Group, (diseñadores y programadores de normas de compresión de audio y video) trabajando bajo la dirección de International Standards



*Reproductores portátiles del formato MP3 de música, que ya incorporan incluso muchos teléfonos móviles, para poder escuchar las canciones grabadas.*

## Formato WAV

El formato WAV (Waveform Audio File) es un formato de archivo originario de Microsoft Windows 3.1, que tiene normalmente la extensión Wav. Es el formato para almacenar sonidos más utilizado por los usuarios de Windows y su flexibilidad lo hace muy adecuado para el tratamiento del sonido pues puede ser compri-



Organization (ISO). Se identifican con la extensión MP3. Esta norma fue lanzada el año 1995 a la Internet, existiendo numerosos Web que contienen ficheros MP3 en sus servidores, y actualmente se trabaja en el sucesor que será el MP4 con una compresión de 40 a 1. La calidad de sonido del MP3 y su pequeño tamaño lo hecho muy popular en Internet; su algoritmo se basa en la forma de escuchar que tiene el oído humano (efecto psicoacústico), pues las frecuencias que quedan fuera de la audición no son registradas en el archivo (las mayores de 20 kHz y las menores de 20 Hz). Esto se traduce en archivos mucho más pequeños, sin una pérdida apreciable de la calidad del sonido.

Al usar el formato de compresión MP3 se puede reducir el tamaño de un CD en un factor de 12 a 1, (1 minuto de calidad CD en formato MP3 equivale aproximadamente a 1MB) pero lo más importante es que no pierde calidad de sonido. El uso más corriente que se da a estos formatos es el de almacenamiento de música (en un disco compacto se pueden grabar unos 12 CD's-Audio, aproximadamente unas 160 canciones) por lo que primero se debe grabar el CD en el disco duro del ordenador para luego pasarlo al formato MP3; para esto se debe tener un programa extractor, llamado Ripper, que cumple la función de pasar desde formato CD a Wav y luego se debe tener un encoder, un programa compresor que pasa de formato Wav a MP3. Se debe tener en cuenta la calidad que se desea, ya que a mayor tamaño de archivo mayor calidad. Esto se especifica eligiendo los kbit/s (512, 256, 128, 64, 32, 20, 16), los kHz (48.000, 44.100, 32.000, 24.000, 22.050, 16.000, 11.025, 8.000) y si es estéreo o mono y como es la calidad del Wav que se está trabajando. Para poder escuchar la canción necesitamos en el PC un reproductor MP3, como el Winamp, o disponer de un reproductor portátil.



WINAMP es uno de los programas más populares para PC, reproductor de archivos de audio en formato MP3.

Respecto a la calidad del MP3, puede ser como uno quiera, menor calidad menor tamaño y menos tiempo empleado, pero para obtener un sonido con calidad digital la tasa de transferencia de bits deberá ser de 128

Kbit/s con una frecuencia de muestreo de 44.100 kHz. Esto hay que tenerlo en cuenta a la hora de crear el MP3, pero por lo general es la opción que viene como estándar en todas las aplicaciones compresoras.

## Formato VQF

El formato VQF (Transform-domain Weighted interleaved Vector Quantization), es un nuevo formato de compresión de audio desarrollado por la empresa japonesa Yamaha, tiene características similares al MP3, pero la compresión y calidad es mejor, se habla de un 30-35% menos de tamaño. Un VQF a 96 kbit/s es tan bueno como un MP3 de 256 kbit/s, pero ocupa  $\frac{1}{4}$  de su tamaño. El formato no está aún muy difundido en comparación con el MP3, pero se cree que es cuestión de tiempo para que los usuarios comparen las ventajas del archivo y su popularidad suba.

Todas estas ventajas se convierten en desventajas al pensar en el hardware, pues necesita mayor cantidad de RAM (mínimo de 16 MB, y se recomienda tener 32 MB) y de velocidad (mínimo Pentium II). Otra desventaja es que los archivos no se pueden llevar a un formato Wav aunque se trabaja en esto.

## Formato Real Audio

Este es el formato más usado en Internet por su capacidad de reproducción en tiempo real, lo que significa que mientras se descarga el archivo se escucha el sonido y cuando se termina de bajar ya ha sido reproducido en su totalidad (*streaming*). No hay que esperar a descargarlo completamente para reproducirlo.

Este formato fue desarrollado por la empresa Real Networks <http://www.real.com/>, que ha puesto a disposición de los usuarios el software para recibir y enviar en tiempo real (tanto vídeo como sonido), que puede descargarse de la red gratuitamente o pagando.

El problema más grave que tiene este formato es que puede entrecortarse la reproducción del audio cuando hay interrupción en la señal de datos, lo que puede ocurrir cuando se utiliza un módem muy lento o hay mucho tráfico en la red y llega a congestionarla. Real Network ha desarrollado una mejora en su formato (Real SystemG2), que mejora la calidad del audio. El problema surge en el almacenamiento pues produce archivos demasiado grandes sobre todo para ser enviados por correo electrónico.



El programa Real Player permite tanto escuchar canciones como ver vídeos desde Internet, siempre que se tenga una conexión de alta velocidad como puede ser una línea ADSL o RDSI, sin la necesidad de almacenar nada en el disco duro. Es el más usado por su rapidez, y el sonido se ve mejorado de calidad en las últimas versiones.

Algunos otros formatos que también se utilizan, son:

- **.aiff (Audio Interchange File Format).** Formato de sonido muy simple y popular en Internet, es un formato originario para Macintosh parecido al Wav por su tamaño, pero también puede ser usado en otras plataformas.
- **.voc.** Son similares a los archivos Wav, la diferencia es que traen marcadores de sincronización especialmente para ser usados con imágenes, videos u otros sonidos en aplicaciones multimedia.
- **.mod.** Es la extensión que se aplica a un tipo de fichero que utiliza muestras de instrumentos digitalizados para crear composiciones en secuenciadores especiales llamados trackers; se comenzó usando solo 4 pistas y actualmente llegan a 32, la calidad sonora es buena y se sigue mejorando. Puede tomar como muestra cualquier sonido digital, el tamaño de los archivos es pequeño y para la reproducción se requiere software especializado.

En resumen, la gran aportación que hacen todos estos formatos aplicados al sonido digital es la de su tratamiento igual que si de cualquier otro tipo de información se tratase, toda vez que se están manejando una secuencia de ceros y unos (0,1), manteniendo el nivel de calidad deseado. A pesar de todo esto, hay que tener en cuenta que nosotros apreciamos el sonido en formato analógico, por lo que la parte final de reproducción (amplificadores, altavoces, etc.) tienen su importancia y no se debe descuidar si queremos tener una audición que satisfaga nuestros gustos musicales, que si son muy exigentes requerirán disponer de un sistema completo Dolby Digital para apreciar en todo su esplendor las grabaciones.

## EL DOLBY DIGITAL

Para los que sean muy exigentes y no se conformen con un equipo de sonido normal existe un sistema denominado Dolby Digital que consigue una calidad extraordinaria, siempre que la instalación esté adecuadamente realizada y todos los equipos de la cadena cumplan las especificaciones que se exigen. Este sistema es válido para audiciones en casa (se puede conectar al televisor y al DVD) y también en salas de cine.

El Dolby Digital es un sistema denominado de 5,1 canales debido a que proporciona 5 canales independientes (izquierdo, derecho, central, surround izquierdo y surround derecho); todos ellos reproducen una gama de 20 a 20.000 Hz, frente al Dolby Surround cuya gama va de los 100 a los 7.000 Hz. Además, el sistema puede proporcionar un canal subwoofer (3-120 Hz) opcional e independiente. A pesar de que los cinco canales proporcionan un ancho total de banda que abarca todo el espectro audible, se añade un canal para los efectos sonoros de Baja Frecuencia para aquellos aficionados que exigen sonidos graves particularmente poderosos, que casi se perciben físicamente. Como los graves son muy difíciles de conseguir - requiere un altavoz muy grande-, pero tiene la ventaja de que no es nada directivo, lo que se hace entonces es poner un altavoz solo para graves en cualquier sitio de la sala de audición. Además, el sistema Dolby Digital aumenta la separación entre canales y la capacidad de que sonidos individualizados lleguen desde múltiples direcciones al oyente.