

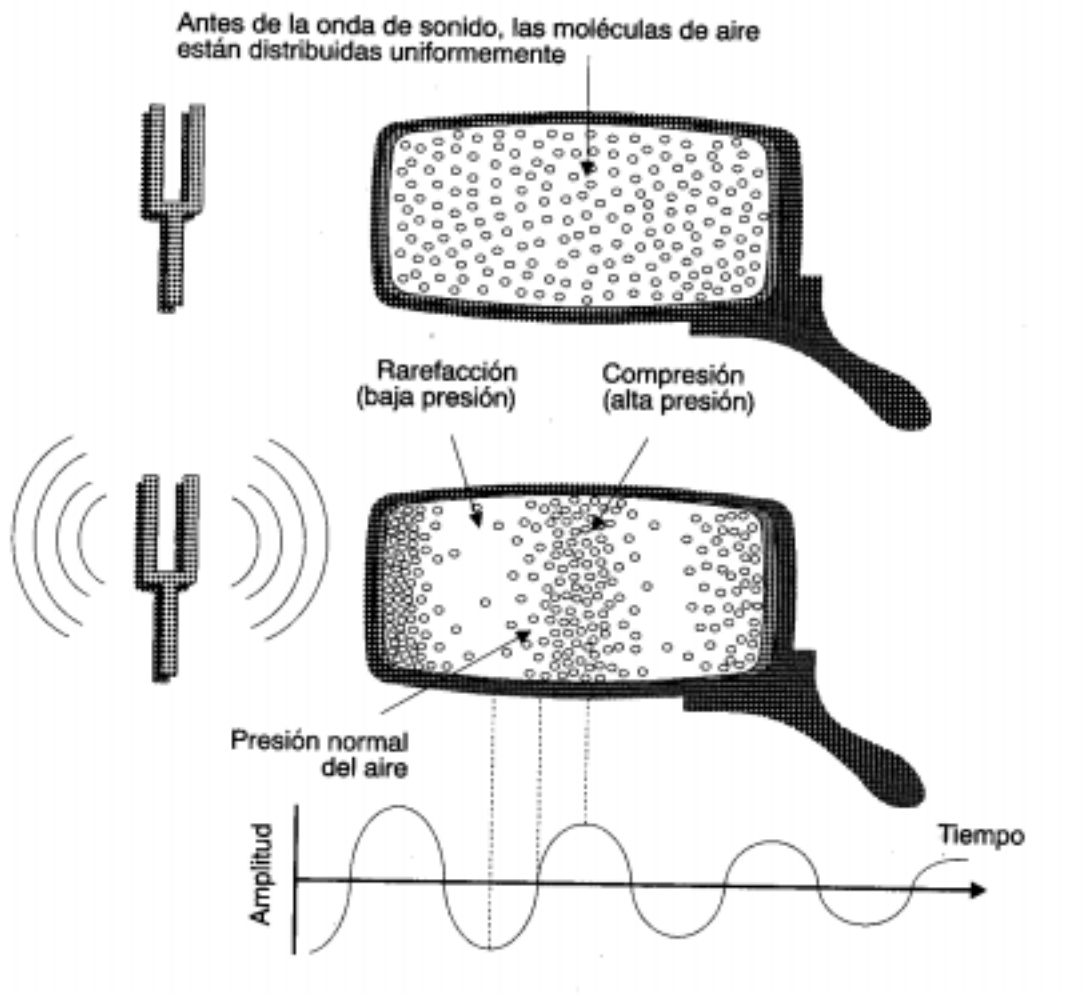
Comunicación multimedial

Unidad 3

Sonido y Audio Digital

La naturaleza del sonido

El sonido es una vibración que se propaga a través del aire, gracias a que las moléculas del aire transmiten la vibración hasta que llega a nuestros oídos. Se aplican los mismos principios que cuando se lanza una piedra a un estanque: la perturbación de la piedra provoca que el agua se agite en todas las direcciones hasta que la amplitud de las ondas es tan pequeña, que dejan de percibirse. La Figura muestra las vibraciones físicas de un diapasón que ha sido golpeado. Las vibraciones del diapasón fuerzan a las moléculas de aire a agruparse en regiones de mayor y menor densidad, dando lugar a que la presión del aire aumente o disminuya instantáneamente. El diapasón es un excelente ejemplo de fuente de sonido por dos razones: la primera es que puede observarse el movimiento de vaivén de sus brazos mientras se escuchan los resultados de esta vibración; la segunda es que el diapasón vibra a *una frecuencia* (vibraciones por segundo) constante hasta que toda su energía se ha disipado en forma de sonido. Una perturbación que viaja a través del aire se denomina *onda* y la forma que adopta esta se conoce como *forma de onda*.



Características de una onda sencilla

La forma de onda del diapasón es la más sencilla de las formas de onda, denominada *onda Sinusoidal*.

La frecuencia de sonido se mide en unas unidades denominadas Hertzios (Hz). Un sonido que vibra una vez por segundo tiene una frecuencia de 1 Hz. Las frecuencias se escriben normalmente en kilohertzios (kHz), unidad que representa 1.000 Hz. Una persona joven con unos oídos saludables puede oír sonidos que estén en el rango de los 20 Hz a los 20.000 Hz (20 kHz).

En la Figura se observa cómo la amplitud de la onda disminuye a medida que el sonido se aleja de su fuente, extendiéndose en todas las direcciones. (La figura exagera la rapidez con la que la onda disminuye en el aire).

El oído y un micrófono incorporado en la tarjeta de sonido se comportan de manera similar. Ambos transforman pequeñas variaciones en la presión del aire en señal eléctrica que puede ser comprendida y almacenada por sus respectivos "cerebros" (ya sea el humano o la CPU de la computadora). Esta señal eléctrica puede ya ser guardada, manipulada o reproducida mediante los medios electrónicos adecuados.

Amplitud

La medida de la amplitud de una onda es importante porque informa de la fuerza, o cantidad de energía, de una onda, que se traduce en la intensidad de lo que oímos, su unidad de medida es el decibelio. Un decibelio, abreviado como dB, es una unidad de medida de la fuerza de la señal y es útil en la comparación de la intensidad de dos sonidos. La sensibilidad del oído humano es extraordinaria, con un rango dinámico o variación en intensidad muy amplio. La mayoría de los oídos humanos pueden capturar el sonido del murmullo de una hoja y, después de haberse sometido a ruidos explosivos como los de un avión, siguen funcionando. Lo que es sorprendente es que la fuerza de la explosión en un avión es al menos 10 millones de veces mayor que el murmullo que una hoja produce con el viento.

El oído necesita un porcentaje elevado de variaciones en la fuerza de un sonido para detectar un cambio en la intensidad percibida, lo que significa que la sensibilidad del oído a la fuerza del sonido es logarítmica, de manera que el decibelio, unidad de medida logarítmica, es la elección más adecuada para medir la fuerza del sonido. El aspecto práctico de la amplitud es que un incremento de sólo 3 dB duplica la intensidad de un sonido. Por ejemplo, un sonido con 86 dB tiene, el doble de fuerza que un sonido con 83 dB y cuatro veces más que un sonido con 80 dB. Desde la perspectiva de nuestra percepción de la intensidad, un incremento de 3 dB, que da lugar a que se duplique la fuerza, provoca que el sonido se perciba sólo ligeramente más alto. Es necesario un aumento en 10 dB para que nuestros oídos perciban un sonido con el doble de intensidad.

Rango dinámico

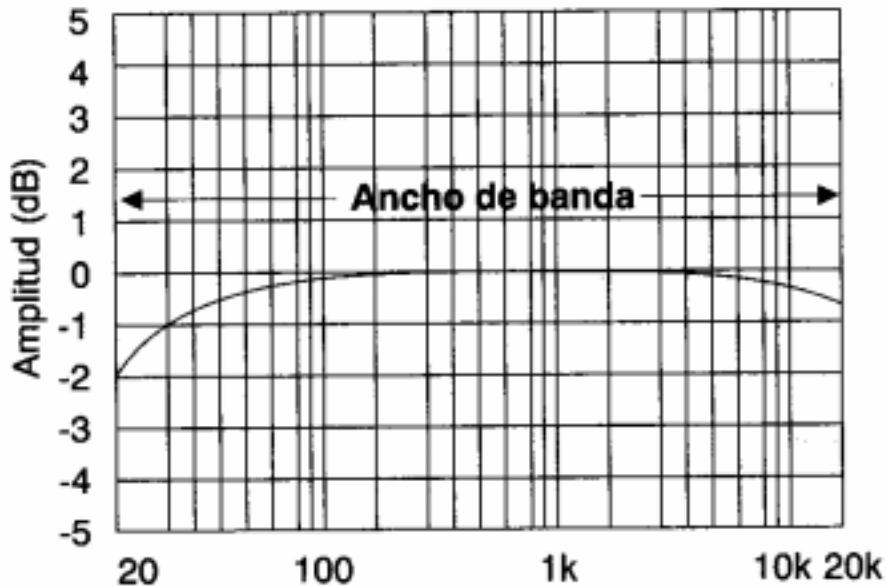
La calidad de los sonidos musicales grabados no es demasiado importante, ya que nunca son comparables a los reales. La razón principal es que el equipo estéreo no puede duplicar el rango dinámico completo de una orquesta o de un concierto de rock. Una orquesta puede alcanzar los 110 dB en su clímax y en el punto más suave bajar hasta los 30 dB, dando lugar a un rango dinámico de 80 dB. Este rango es superior al rango dinámico de un sistema estéreo típico y, de hecho, superior a la capacidad de grabación de medios tales como un disco de vinilo y una cinta de audio.

Ancho de banda

Profundizamos ahora en aspectos prácticos, como el rango de frecuencia con el que es capaz de trabajar un reproductor CD, nuestro oído o nuestra voz. El ancho de banda es muy importante para disfrutar de la música (como manifiestan las quejas de sonido "de lata" de una radio de bolsillo) y es un criterio básico a la hora de seleccionar un equipo de audio para utilizar con la tarjeta de sonido. Por ejemplo, el ancho de banda teórico de la radio FM es aproximadamente tres veces el ancho de banda de la radio AM, por lo que la FM será capaz de reproducir frecuencias que no entran dentro del campo de trabajo de la AM.

Nota: A menudo el ancho de banda se simboliza mediante un único número cuando la frecuencia baja está bastante próxima a cero. Por ejemplo, el ancho de banda de una voz femenina se sitúa en torno a los 9 kHz, aunque realmente puede estar en el rango que va desde los 200 Hz hasta los 9 kHz.

Existe una medida estándar para definir el ancho de banda: el rango de frecuencias sobre el que la amplitud de la señal no difiere del promedio en más de 3 dB, es decir la diferencia de las frecuencias en la que se produce una caída de 3 dB, ya es el punto donde su amplitud cayó a la mitad, y éste es el mínimo cambio en la fuerza de la señal que puede ser percibido como un cambio real en la intensidad por la mayoría de los oídos.



Ancho de banda del sistema

Es importante tener en cuenta que el ancho de banda de un equipo de sonido depende del enlace más débil del canal, que normalmente no es la tarjeta de sonido. La calidad del sonido producido por la computadora refleja el esfuerzo de muchas componentes, y la salida no será mejor que la interpretación del miembro menos capacitado de un grupo. En el caso del sistema de sonido de la computadora, una señal debe pasar por muchas fases de transformación de audio y por diferentes dispositivos. Por ejemplo, consideremos el sonido grabado mediante un micrófono y que luego es reproducido. La tarjeta de sonido transforma el sonido recogido del micrófono en una señal eléctrica que, posteriormente, se transforma en audio digital y se almacena en disco. El audio digital del disco es transformado de nuevo en una señal eléctrica y reproducido a través de los cascos o de los altavoces. El ancho de banda efectivo del sistema de sonido está limitado por el dispositivo con el ancho de banda más estrecho de todos los dispositivos que procesan el sonido.

El enlace más débil en grabación suele ser el micrófono, que tiene probablemente un ancho de banda aproximadamente de 12 kHz.

Ruido

Del mismo modo que perturban los ruidos y ecos en una habitación, también puede generarse ruido y distorsión en la tarjeta de sonido, en los altavoces y en el micrófono. El ruido (sonidos aleatorios que transforman y enmascaran el sonido deseado) se mide también en decibelios. Dado que es tan poco probable disponer de un entorno de audio digital en perfecto silencio, lo que interesa realmente es saber la cantidad de ruido en relación con la señal que se introduce en el equipo de sonido, especialmente en la tarjeta de sonido. La fuerza de la música, del habla o de cualquier otro sonido, comparada con la fuerza promedio del ruido, se conoce como relación señal/ruido. A medida que aumenta la relación s/n, es mejor el trabajo realizado en grabación. La relación señal ruido de una tarjeta digital sencilla es del orden nada despreciable de 85 dB. Esto significa que la fuerza de la señal es 85 dB mayor que la fuerza del ruido. Una relación de 70 dB se considera válida para propósitos musicales y una relación de 65 dB s/n está en el límite de aceptación.

Grabación y reproducción de audio: Bases del audio digital

Antes de que la computadora pueda grabar, manipular y reproducir sonido, debe transformarse el sonido de una forma analógica audible a una forma digital aceptable por la computadora, mediante un proceso denominado conversión analógica - digital (ADC). Una vez que los datos de sonido se han almacenado como bytes en la computadora, puede hacerse uso de la potencia de la CPU de la computadora para transformar este sonido de miles de modos. Con el software adecuado es posible, por ejemplo, añadir reverberación o eco a la música o a la voz. Pueden eliminarse trozos de sonido grabado. Pueden mezclarse archivos de sonido, ajustarse el tono de la voz de manera que no pueda reconocerse y muchas cosas más. Finalmente, cuando se está dispuesto a escuchar el resultado, el proceso de conversión digital-analógica (DAC) transforma de nuevo los bytes de sonido a una señal eléctrica analógica que emiten los altavoces.

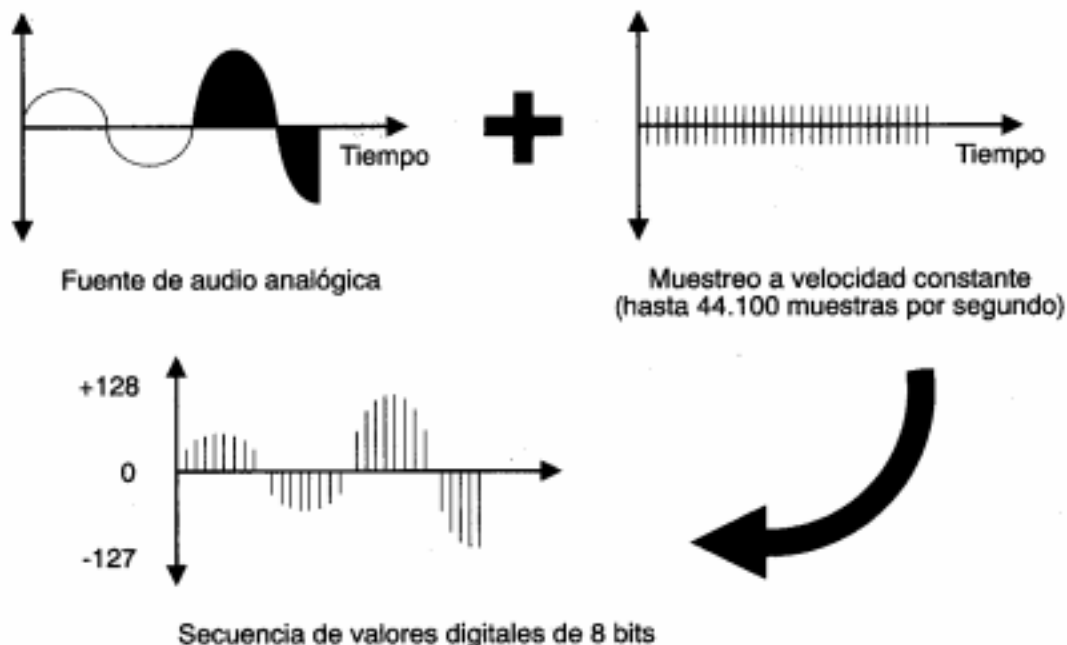
Muestreo: Conversión analógica-digital y viceversa

Comenzaremos con la captura del sonido haciendo uso del micrófono. Cuando las ondas de sonido llegan al micrófono, el movimiento mecánico se traduce en una señal eléctrica. Esta señal se denomina señal analógica porque es una señal continua en el tiempo, análoga al sonido original.

CONVERSION ANALOGICA-DIGITAL (ADC): Dada una señal analógica, se van tomando valores discretos de su amplitud a intervalos de tiempo pequeños, evidentemente será más fiable la reproducción cuantas más muestras por segundo se tomen. A estos valores obtenidos se les asigna un valor digital que el computadora puede entender y procesar como se requiera.

A cada muestra obtenida se le asigna un equivalente binario ya que es en este sistema en el que trabajan las computadoras, su unidad de información es el bit. Un bit solo puede tomar dos posibles valores "1" o "0", es lógico pensar que será necesario ampliar esta unidad de información para así poder asignar a cada valor de muestra tomada un equivalente binario. Por esta razón y dependiendo de la fidelidad con que queramos trabajar podemos utilizar palabras de 8 o 16 bits pudiendo obtener así 256 ó 65536 combinaciones distintas y obtener mayor resolución.

CONVERSION DIGITAL-ANALOGICA (DAC): El proceso inverso es mucho menos complejo ya que solo se trata de ir poniendo los valores de las muestras en el mismo orden que fueron tomados y unos filtros electrónicos se encargan de convertir esa señal resultante de valores discretos en una señal analógica.



Velocidad y tamaño de muestra

La fidelidad terminología empleada por los entusiastas del audio para expresar la exactitud en la réplica de la música original del sonido audio digital depende de la selección de la correcta frecuencia de muestreo y del correcto tamaño de muestra, siendo este último el número de bytes utilizados para almacenar cada muestra.

FRECUENCIA DE MUESTRA: La frecuencia de muestra (también denominada frecuencia de muestreo) debe ser lo suficientemente alta para que los sonidos de alta frecuencia, como el sonido del cristal de una copa de vino o el del arco de un violín, puedan recogerse con precisión. Según el teorema de Nyquist, es posible repetir con exactitud una forma de onda si la frecuencia de muestreo es como mínimo el doble de la frecuencia de la componente de mayor frecuencia. La frecuencia más alta que puede percibir el oído humano está cercana a los 20 kHz, de modo que la frecuencia de muestreo de 44.1 kHz de las tarjetas de sonido es más que suficiente. Este valor es el utilizado hoy en día por los reproductores de audio CD.

Los archivos de audio digital pueden grabarse seleccionando la frecuencia de muestreo. A medida que aumenta la frecuencia de muestreo, aumenta la calidad del sonido. Por ejemplo, una velocidad de 6.000 Hz (6.000 muestras por segundo) es buena para una voz masculina típica, pero no lo es para una voz femenina típica, que tiene componentes con una frecuencia más alta. Una frecuencia de muestreo de 8.000 Hz proporciona una grabación de la voz femenina de mayor calidad. La siguiente tabla ofrece una lista de algunas tarjetas de sonido Sound Blaster y de sus frecuencias de muestreo:

Tarjeta Grabación - Reproducción

Sound Blaster 16	44.100 (mono o estéreo)	44.100(mono o estéreo)
Sound Blaster Pro	22.050 (estéreo)	22.050(estéreo) o 44.100(mono)
Sound Blaster 2.0	15.000 (mono)	44.100(mono)
Sound Blaster 1.0,1.5	13.000 (mono)	23.000 (mono)

La Sound Blaster 16 puede grabar en estéreo, grabando hasta 44.100 muestras por segundo, con un canal izquierdo y otro derecho que producen una frecuencia de muestreo combinada de 88.200 muestras por segundo. Las tarjetas Sound Blaster Pro y la Sound Blaster 16 son capaces también de trabajar en estéreo con una velocidad máxima de reproducción de 22.050. Ambas tarjetas, la Sound Blaster y la Sound Blaster Pro, toman muestras de sonido de 8 bits (1 byte); cada medida consume 1 byte de almacenamiento de la memoria de la computadora o del disco. La Sound Blaster 16 maneja muestras de 16 bits (2 bytes), emitiendo voz y música con una fidelidad equivalente a los reproductores CD actuales.

Existen varias razones para no utilizar las frecuencias de muestreo más altas. En primer lugar, las frecuencias de muestreo altas necesitan gran capacidad de almacenamiento.

TAMAÑO DE MUESTRA:

El tamaño de muestra es la otra componente de mayor influencia en la fidelidad del audio digital. Las tarjetas de sonido de 16 bits ofrecen la posibilidad de elegir entre un tamaño de muestra de audio digital de 8 bits (1 byte) o de 16 bits (2 bytes).

El tamaño de muestra controla el rango dinámico que puede grabarse. Por ejemplo, las muestras de 8 bits limitan el rango dinámico a 256 pasos (rango de 50 dB). Por el contrario, una muestra de 16 bits tiene un rango dinámico de 65.536 pasos (rango de 90 dB) una mejora sustancial. El oído humano percibe todo un mundo de diferencias entre estos dos tamaños de muestra. Los oídos son más sensibles a la detección de diferencias en el tono que en la intensidad, pero son aún más sensibles a la fuerza del sonido. Los oídos humanos, que están acostumbrados a detectar sonidos con variaciones de varios órdenes de magnitud en la fuerza, perciben el sonido de 8 bits en un tono apagado o desafinado si se compara con el sonido de audio digital de 16 bits.

COMPROMISOS EN EL MUESTREO Se podría asumir que todo lo que hay que hacer para obtener buen sonido es grabar a la velocidad límite de 44,1 kHz con muestras de 16 bits (2 bytes). El único problema que aparece si se graba en estéreo, tomando muestras simultáneamente en los canales izquierdo y derecho a 44,1 kHz, una muestra de sonido de un minuto necesita un espacio para almacenarse de 10,58MB.

Lo aconsejable es usar la frecuencia de muestreo más baja posible. Por ejemplo, supongamos que planeamos grabar una conversación telefónica. El ancho de banda de un teléfono es de sólo 3 kHz. De acuerdo con el teorema de Nyquist, la grabación será acertada si la frecuencia de muestreo es de 6 kHz o mayor.

Cuando se elige la frecuencia de muestreo, también hay que considerar el ancho de banda de todo el sistema. Por ejemplo, no existe ningún problema en la grabación de audio digital a 44,1 kHz si el micrófono utilizado funciona a 12 kHz y la fuente de sonido es una voz masculina grave que no supera los 7 kHz.

Compresión de audio digital

Como hemos visto anteriormente, cada muestra tomada es almacenada en el ordenador ocupando uno o dos bytes de memoria dependiendo de si se ha tomado una resolución de 8 o 16 bits. Esto conlleva emplear grandes espacios de disco para almacenar estos ficheros de sonido. Se han desarrollado muchos formatos de archivos comprimidos que permiten realizar grabaciones de alta calidad sin necesidad de utilizar tanto espacio de disco. Estas técnicas suelen incluir un método para codificar secuencias largas de bytes repetidos. Evidentemente existe una contrapartida y es que una vez comprimido el archivo de sonido, este no puede ser editado para su modificación.

EL SONIDO EN MULTIMEDIA

Tal vez sólo antecedido por los gráficos, el sonido es una herramienta primordial en el mundo multimedia; en especial en Internet, mientras navegamos, encontramos en mayor cantidad sitios con formatos de audio que nos hacen más placentera nuestra estancia en dicho sitio. Empero dichos formatos (así como los formatos de animación y algunos de video) todavía no están disponibles a un nivel de "profesionales" ante nosotros. Actualmente el ver imágenes, videos o animaciones están solo completos, como una forma de absorber la información, con los sonidos. Gracias a ello podemos contar con un formato más adecuado, con resultados potencialmente mayores, gracias a que el mensaje que se quiere transmitir es captado con mayor facilidad; así como una comunicación más efectiva y amena. A pesar de la importancia del audio (o tal vez gracias a esta) no existe un formato universal que los usuarios y los desarrolladores de multimedia usen más; este varía dependiendo de el usos que se le de a dicho formato. La mayor dificultad que enfrentan es su tamaño, ya que la transferencia de audio es mucho más tardada que la transferencia, por ejemplo, de todo el texto de una página promedio. Anteriormente, en el artículo "La Música Electrónica", hable de uno de los formatos de audio más comúnmente usados (por su compacto tamaño) el MIDI, sin embargo, la desventaja de este es la falta de calidad en su reproducción, y que este no puede albergar mas que sonidos sintetizados, y no la versatilidad de sonidos "reales", como la voz, que tiene la música. En este artículo quiero mostrarles algunos formatos de audio que podemos utilizar los que no poseemos un estudio profesional de grabación.

MICROSOFT WAVE (*.WAV) Tal vez el más usado por los usuarios de Windows, este formato de Microsoft WAV (Waveform Audio File) es un formato muy flexible que puede ser comprimido y grabado en diferentes tamaños y rangos. Aunque los archivos WAV pueden archivar excelente calidad de audio, comparable con la calidad de CD (44 kHz a 16 Bits en estéreo, aproximadamente) el tamaño necesario para crear dicha calidad es extremadamente grande e ineficiente (en especial para usarlo en Internet) así que la opción opuesta que sería grabar a 4 Bits y kHz (los más bajo que puede grabar este formato) produce una calidad bastante baja; el rango bajo de kHz produce estática, ruido e incluso cortes en la reproducción del Formato. Por esta razón dichos formatos solo se usan como muestras, y solo en casos muy especiales existe alguna melodía completa grabada bajo este formato. Por otro lado la flexibilidad y compatibilidad de los MS WAV permiten que se puedan convertir a varios formatos por medio de editores o codificadores.

MACINTOSH (*.AIFF) AIFF (Audio Interchange File Format) es uno de los formatos de audio más simples usados por las aplicaciones de Macintosh. Este tipo de archivos desarrollados por Apple Computer para usar audio data

(datos de audio) es muy semejante al WAV de Microsoft; empero los formatos AIFF sólo pueden ser leídos por computadoras Macintosh, por lo que en ocasiones es recomendable , cuando se desarrolle algún sitio en Internet, usar ambos formatos, para que el usuario pueda acceder, no importando la plataforma sobre la cual navegue. Los Formatos AIFF también pueden grabarse en un rango amplio de kHz, pero tienen el mismo problema de tamaño que los WAV.

REAL AUDIO Este es el formato más usado en Internet por su capacidad de reproducir en tiempo real (real time) lo que significa que mientras es descargado el archivo, este es tocado, así que cuando termina de "bajar", ya fue completamente reproducido. Aunque la calidad no es muy alta, el ahorro en tiempo lo convierte en uno de los formatos favoritos para los creadores de páginas en la Web. Este formato fue desarrollado por RealNetworks. Con su locación el Seattle, está empresa a desarrollado e introducido al mercado software y servicios que permitan al usuario ejecutar, recibir o enviar real time media (tanto en video como en audio) Esta empresa está dentro de las 100 empresas más influyentes en el ámbito computacional; el crecimiento y desarrollo a través del Internet, a puesto al RA (Real Audio) a la altura de formatos tan usados como el WAV o incluso el MIDI. Empero, la gran ventaja de este formato es también su lado flaco, ya que a pesar de que normalmente funciona de manera óptima, la reproducción se ve "cortada" cuando se encuentran interrupciones en la señal de datos; esto usualmente pasa cuando el usuario está conectado a un modem lento o existe mucho tráfico en la red; el resultado en una reproducción con pausas y poco satisfactoria. RealNetworks ha desarrollado un nuevo sistema que pronto va a ganar un espacio en la reproducción de audio; es el RealSystem G2, que incrementando al respuesta de frecuencia de audio a un 80% aprox. En modems desde 28.8 Kbps logra una calidad superior de audio Ahora, este formato a pesar de que puede reproducir en diferentes rangos, dependiendo de la configuración del usuario, guarda sus archivos en tamaños mucho más grandes que el RA tradicional, lo que puede causar problemas al bajar el archivo, tal vez no tan grandes como los WAVs o AIFF, pero si considerables para almacenamiento y sobre todo envío (por mail). Este reproductor también puede tocar archivos WAV y los formatos antiguos de RA. Y puede ser usado por usuarios de PC o MAC.

MPEG AUDIO: MPEG-1 Layer-3 (*.MP3) El formato de Audio desarrollado por MPEG (Moving Picture Experts Group) trabajando bajo la dirección de International Standards Organization (ISO) y the International Electro-Technical Commission (IEC) ha logrado archivos de sonido compresos en ratios de 12-14:1 con calidad de CD estéreo. La historia del MP3 empieza en el Institut Integrierte Schaltungen (IIS) en Alemania en 1987; el IIS comenzó a trabajar en la codificación de audio EUREKA, proyecto para la transmisión Digital de Audio (Digital Audio Broadcasting o DAB) En cooperación con la Universidad de Erlangen el ISS finalmente encontró un algoritmo ahora llamado ISO-MPEG Audio Layer-3. Layer-3 es tan eficiente que puede reducir los tamaños de los archivos dramáticamente a través de sus algoritmos de reducción de datos. Sin la reducción de datos, las señales típicas consisten en archivos de 16 Bits a un rango de 44 kHz (lo usado por los CD) por lo que representar un solo segundo de música en estéreo y con calidad de Compact Disc tenemos un archivo de 1.400 Kbit aproximadamente. Al usar el formato de codificación MPEG se puede reducir la pista desde el CD a un factor de 12 a 1, pero lo más importante, sin perder calidad de sonido. Factores de reducción de incluso 24 mantienen una calidad bastante aceptable aún, ya que la codificación tiene, dentro de su algoritmo, técnicas que están basadas en la percepción de ondas sonoras de oído humano.

Existen tres formatos de MPEG:

Factor de compresión de 4:1 del Layer 1 (correspondientes a 384 kbps de señal estéreo).

Factor de compresión de 6-8:1 del Layer 2 (correspondientes a 256 kbps de señal estéreo).

Factor de compresión de 10-14:1 del Layer 3 (correspondientes a 128 kbps de señal estéreo).

El formato MP3, no solo conserva una excelente calidad, sino que amplía los efectos estéreo, limitando un poco el ancho de banda. Para darnos una idea del desempeño del MP3 incluyo la siguiente tabla:

Sonido	Ancho de banda	de Modo	Kbps	Radio de Reducción
Teléfono	2.5 kHz	mono	8	96:1
Señal de Onda Corta	4.5 kHz	mono	16	48:1
Amplitud Modulada	7.5 kHz	mono	32	24:1
Frecuencia Modulada	11 kHz	estéreo	56...64	26...24:1
MP3	10..15 kHz	estéreo	96...128 kbps	12...10:1
CD	15 kHz	estéreo	112..128 kbps	14..12:1

Estos son los formatos más usados para reproducir audio si requieren más información pueden consultar las siguientes páginas:

<http://www.microsoft.com>

<http://www.real.com>

<http://mp3.com>

MIDI

El Musical Instrument Digital Interface (MIDI) es el protocolo de comunicaciones estándar utilizado para intercambiar datos entre sintetizadores, ordenadores, software, procesadores de efectos y otros dispositivos MIDI. Los archivos MIDI son significativamente diferentes de los archivos de audio digital como los VOC o los WAV. Los archivos de audio digital contienen sonido real grabado de forma digital a partir de la captura de miles de muestras por segundo. La música MIDI, por otro lado, contiene solo instrucciones sobre la manera de tocar un instrumento. Los archivos de audio digital pueden necesitar millones de bits para almacenar unos minutos de música, mientras que los MIDI con unos pocos miles de bits pueden reproducir música durante horas.

Un archivo MIDI es completamente editable, pudiéndose manipular el instrumento que se toca, la duración de una nota, la intensidad del sonido, etc...