
Conceptos básicos de sonido digital

PID_00194466

Cinto Niqui i Espinosa



Cinto Niqui i Espinosa

Profesor asociado del Departamento de Comunicación Audiovisual y Publicidad de la Facultad de Ciencias de la Comunicación de la Universidad Autónoma de Barcelona desde 1992. Profesor de la Escuela Superior de Cinematografía y Audiovisuales de Cataluña, adscrita a la Universidad de Barcelona, desde 1994. Consultor de los estudios de Comunicación Audiovisual de la Universitat Oberta de Catalunya desde el 2002. Compagina la docencia universitaria con el trabajo periodístico. Es director y presentador del programa *L'altra ràdio*, de Radio Nacional de España en Cataluña. Este espacio está especializado en la cultura audiovisual y se emite desde el año 1980. Tiene los títulos de ingeniero técnico de Telecomunicaciones en la especialidad de Imagen y Sonido por la Escuela de Ingeniería Técnica de Telecomunicación La Salle Bonanova; máster de Nuevas tecnologías: comunicación audiovisual y publicitaria por la Universidad Autónoma de Barcelona; licenciado en Filología Catalana por la Universidad de Barcelona, y máster de Escritura para la televisión por la Universidad Autónoma de Barcelona.



Los textos e imágenes publicados en esta obra están sujetos –excepto que se indique lo contrario– a una licencia de Reconocimiento-NoComercial-SinObraDerivada (BY-NC-ND) v.3.0 España de Creative Commons. Podéis copiarlos, distribuirlos y transmitirlos públicamente siempre que citéis el autor y la fuente (FUOC. Fundació per la Universitat Oberta de Catalunya), no hagáis de ellos un uso comercial y ni obra derivada. La licencia completa se puede consultar en <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

Índice

Objetivos.....	5
1. La digitalización del sonido.....	7
2. Conversor analógico/digital.....	9
2.1. Filtrado	9
2.2. Muestreo	9
2.3. Cuantificación y codificación en binario	10
2.4. Multiplexación	12
2.5. Detección y corrección de errores	12
2.6. La señal digital electrónica	14
3. Conversor digital/analógico.....	16
4. Formatos de audio.....	19
4.1. Formatos de compresión de audio	20
4.2. Formatos de distribución musical	22

Objetivos

1. Repasar los conceptos básicos de la digitalización de las señales analógicas.
2. Recordar algunas de sus características más importantes.
3. Entender cómo se convierte una señal digital en analógica.

1. La digitalización del sonido

Resumen

La digitalización del sonido ha llevado la creatividad sonora a los hogares. Cualquier ordenador multimedia actual permite a cualquier persona experimentar la grabación, la edición y el procesamiento del sonido.

El audio digital es más inmune a las degradaciones –distorsiones– y a los ruidos que el analógico porque, aunque los pulsos electrónicos se deformen, se pueden reconstruir y regenerar en cada etapa del proceso. Con el paso del tiempo, el sonido digital siempre conserva las características originales. No pierde brillo ni volumen.

La calidad del audio digital solo depende de la frecuencia de muestreo y del número de bits empleados. La calidad solo depende de la codificación y la decodificación de la señal.

Los aparatos digitales tienen una dinámica más buena que los analógicos porque los soportes no tienen ruido de fondo. Por otro lado, la separación entre canales –diafonía– es muy elevada.

Gracias a los sistemas de compresión, el sonido digital de buena calidad ocupa cada vez menos espacio. Además, se pueden hacer copias de un material original sin perder generaciones. En el fondo no se hacen copias, sino originales nuevos.

El acceso a la información es rápido y preciso. Un único aparato puede integrar el resto y hacerlo todo. Es lo que se conoce con el nombre de **estaciones de trabajo de audio digital (DAW¹)**. Sistemas como el Pro Tools HD o el Cool Edit Pro 2 lo son.

⁽¹⁾Sigla de *digital audio workstation*.

La información digitalizada, guardada en la memoria de un ordenador, permite trabajar en red. Se pueden enviar archivos a un ordenador que esté al lado mismo o en las antípodas. En el sistema Pro Tools, la transmisión de archivos de sonido es la función Digi Delivery.

El trabajo en red también posibilita trabajar en simultáneo desde ciudades diferentes en una misma sesión de edición. La aplicación DigiStudio de Pro Tools permite interconectar dos editores que se encuentren en estudios diferentes por Internet y trabajar en un entorno virtual común.

Con las redes informáticas también se puede acceder a las bases de sonidos digitales de música o efectos que permiten a los usuarios comprar efectos de sonido o músicas de librería y cargarlos directamente al ordenador, por ejemplo, la base de datos Digi Pro Net, de Pro Tools.

La edición de sonido digital no es destructiva. Los cortes se hacen sobre copias del original de modo que, si hay algún error, siempre se puede recuperar el original.

La gran capacidad de manipulación del sonido facilita la creatividad sonora. Con el **procesamiento digital de la señal (DSP²)** se pueden hacer muchas transformaciones de un sonido real y también se pueden crear sonidos nuevos por síntesis electrónica.

⁽²⁾**DSP** m Sigla inglesa de *procesamiento digital de la señal*.

Hay sistemas que pueden analizar el ruido de fondo de una grabación digitalizada y crear uno igual, pero en contrafase, para eliminarlo. Con los procesadores psicoacústicos de armónicos se puede devolver el brillo a un sonido apagado. Estas dos funciones permiten la reconstrucción sonora de materiales antiguos.

Los ordenadores aplicados a la radiodifusión permiten automatizar la emisión total o parcialmente.

Automatización total o parcial de la emisión

En los ordenadores empleados en las radiofórmulas, la combinatoria del sistema tiene en cuenta que un disco no suene siempre a la misma hora y minuto, o seguido siempre de un mismo tema.

Algunos de los programas diseñados para la radio informan de los anuncios emitidos y gestionan la facturación automáticamente.

2. Conversor analógico/digital

Repasemos brevemente las fases del proceso de digitalización de una señal analógica: filtrado de la señal analógica, muestreo, cuantificación y codificación, multiplexión de los canales, inserción de códigos de detección y corrección de errores y de los valores binarios en la señal electrónica modulada.

2.1. Filtrado

Antes de la digitalización del sonido, se filtra la señal analógica por dos razones:

- 1) Para procesar solo las frecuencias que podemos escuchar, es decir, por debajo de los 20 kilohercios, de forma que el resto no entren en la cadena de la digitalización y no ocupen memoria o ancho de banda innecesario.
- 2) Y porque hay que filtrar todas las frecuencias por encima de la mitad de la de muestreo para evitar la generación de frecuencias indeseables en el margen audible, lo que en inglés se conoce como *filtré antialiasing*.

Por eso, se hace pasar la señal analógica por un filtro pasa bajo con frecuencia de corte de 20 kilohercios. Inmediatamente después, la señal filtrada entra en el circuito de retención y medición (*sample and hold*) para ser digitalizada.

2.2. Muestreo

El muestreo consiste en tomar una muestra de la señal analógica cada cierto tiempo, según lo que marque la frecuencia de muestreo, para hacer una exploración de una señal continua. Después de la exploración, se obtienen una serie de valores instantáneos discretos.

La frecuencia de muestreo tiene que ser como mínimo el doble (2,2) de la frecuencia más alta del audio que se digitalice (criterio de Nyquist).

Si no se hace así, la muestra analizada no se correspondería con la forma de la onda original y surgirían unas frecuencias ajenas al margen auditivo que provocarían distorsión armónica. El efecto de un muestreo insuficiente se conoce en inglés como *aliasing* (*alias frequencies*).

La frecuencia más alta para codificar en audio es de 20 kilohercios; entonces, la frecuencia de muestreo mínima tiene que ser 44 kilohercios. Por ese motivo, los aparatos digitales de primera generación usaban las frecuencias de 44,1 ki-

Lectura recomendada

Encontraréis muchos datos técnicos sobre la digitalización de las señales analógicas y sus parámetros en el temario de la asignatura de la UOC *Fundamentos de tecnología audiovisual*.

lohercios o 48 kilohercios. Para usos de radiodifusión se escogió una frecuencia de 32 kilohercios, dado que en la difusión en frecuencia modulada solo se transmite audio de baja frecuencia hasta 15 kilohercios.

Las frecuencias de muestreo más empleadas

En abril de 1978, la AES (Audio Engineering Society) adoptó el uso de los 44.056 hercios (44,1 kilohercios) como frecuencia estándar de muestreo. En el mismo año, la Asociación de Industrias Electrónicas Japonesas (EIAJ) estableció el formato de cuantificación de 14 bits para usos domésticos.

Resumimos a continuación las frecuencias de muestreo más empleadas:

- 11 kilohercios: ficheros de sonido de los primeros CD-ROM multimedia.
- 32 kilohercios: da una calidad de sonido similar a la de la radio en frecuencia modulada.
- 44,1 kilohercios: empleada en los discos compactos.
- 48 kilohercios: es un valor que se utilizaba mucho en los aparatos de sonido profesional de los noventa.
- 96 y 192 kilohercios: valores que usan los equipos más modernos de gama alta (como Pro Tools HD o Cool Edit Pro 2, entre otros). En 1994, Pioneer Electronics presentó el Wide Band DAT-Recorder, que ya trabajaba a 96 kilohercios y 32 bits.

Cuanto más alta sea la frecuencia de muestreo empleada, más exactitud se obtendrá en la codificación del sonido, la respuesta transitoria será más buena y, lógicamente, la fidelidad sonora también. Pero las frecuencias de muestreo demasiado altas (96 o 192 kilohercios) tienen otro objetivo: permiten manipular repetidamente una misma señal con varios efectos y con la máxima precisión. Cuanto más muestras se tengan de la señal original, de más valores dispondrá el sistema informático para hacer los cálculos de los algoritmos con más exactitud. El aumento de la velocidad de los procesadores, que tratan millones de datos por segundo, ha permitido incrementar las frecuencias de muestreo.

Ejemplo de frecuencias de muestreo

Escuchemos el mismo sonido, siempre digitalizado con una cuantificación de 16 bits por palabra, pero con tres frecuencias de muestreo diferentes:

- La primera: 44,1 kilohercios. El fichero ocupa 126 kilobytes.
- La segunda: 8 kilohercios. El fichero ocupa 23 kilobytes.
- Y la de peor calidad: 2 kilohercios. El fichero ocupa 5,86 kilobytes.

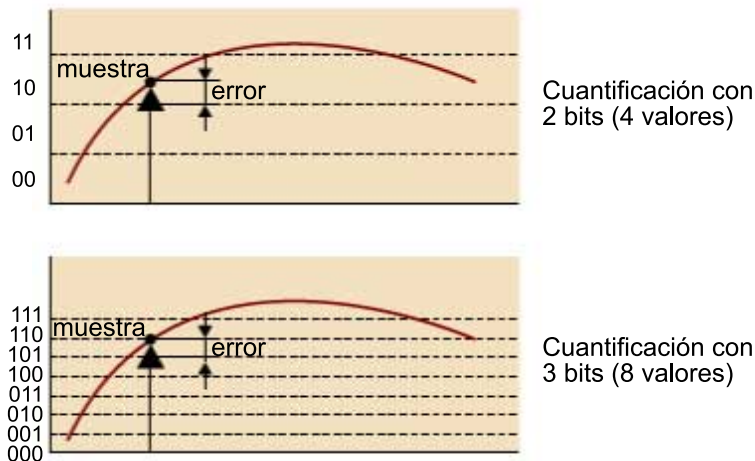
2.3. Cuantificación y codificación en binario

Una vez obtenidas las muestras de una señal, se cuantifican. Es decir, se miden los voltios de cada muestra y el valor del voltaje de cada medición se codifica en forma de palabra binaria. El bit de la izquierda de cada palabra es el más significativo (MSB³) y el de más a la derecha, el que lo es menos (LSB⁴).

⁽³⁾Sigla de *most significant bit*.

⁽⁴⁾Sigla de *less significant bit*.

Para saber el número de niveles de cuantificación que obtenemos con 8 o 10 bits, hacemos 2^8 (256 niveles) o 2^{10} (1.024 niveles). Cuanto más número de bits por palabra, más exacta será la medición del valor de la amplitud de la muestra.



La resolución también determina el valor de la relación señal/ruido (S/N). Aproximadamente, cada bit añadido hace subir este parámetro en unos 6 decibelios.

Número de niveles de cuantificación	Número de bits	Relación señal/ruido (dB) (aproximada)
64	6	37
256	8	49
1.024	10	61
65.536	16	97
1.048.576	20	121
16.777.216	24	145

En cuanto al número de bits por muestra, las cantidades más habituales son las siguientes:

- 8 bits: sonido de los primeros CD-ROM multimedia.
- 16 bits: sonido profesional de los noventa.
- 24 bits: sonido profesional de los noventa de gama alta.
- 32 bits: equipos profesionales más modernos.

Formato de cuantificación para usos domésticos

En 1978, la Asociación de Industrias Electrónicas Japonesas (EIAJ) estableció el formato de cuantificación de 14 bits para usos domésticos.

Cuantos más bits por palabra –o más frecuencia de muestreo–, más espacio de memoria ocupado, mayor debe ser el ancho de banda del canal de transmisión o se tiene que poder trabajar con *bit rates* más elevadas. Recordad que ocho bits consecutivos forman un byte (kB, kilobyte; MB, megabyte).

Si se trabaja con un número de bits bajo, se pueden producir indefiniciones en los valores de las mediciones de las partes más débiles de la señal, que en el momento de la reproducción producirían un efecto parecido a la distorsión armónica. Para evitarlo, a la entrada del conversor de analógico a digital, las señales flojas se superponen a un ruido blanco aleatorio que hace de almohada. Este ruido aleatorio, denominado *dither*, mejora la resolución del proceso, dado que hace subir la señal analógica por encima del nivel del bit menos significativo y la mantiene siempre dentro de los márgenes de cuantificación. Esto reduce la distorsión armónica y la de intermodulación.

Ejemplo de cuantificación

Escuchemos el mismo sonido original siempre digitalizado con una frecuencia de muestreo de 44,1 kilohercios, primero con una cuantificación de 16 bits por palabra. El fichero ocupa 126 kilobytes.

Ahora lo digitalizamos con una cuantificación de 8 bits por palabra. El fichero ocupa, lógicamente, la mitad, 63,3 kilobytes. En este caso concreto, la calidad de sonido no cambia mucho. Por lo tanto, si lo tuviéramos que enviar, nos podríamos ahorrar tiempos de transmisión si cuantificáramos la información a 8 bits.

2.4. Multiplexación

Muchos sistemas que trabajan con información de dos canales, izquierdo y derecho, multiplexan los datos binarios. Eso significa que cada bloque de muestras contiene información del canal derecho y del izquierdo, de forma que los datos de los dos canales podrían pasar por una única línea.

Si multiplexáramos palabras binarias de cuatro bits del canal derecho "dddd" y del izquierdo, "eeee", obtendríamos estas nuevas ráfagas binarias, "ddee" y "ddee".

2.5. Detección y corrección de errores

En audio digital hay dos tipos de errores básicos:

1) Los **errores de ráfaga**, que provocan la pérdida de varias muestras sucesivas. A veces los puede provocar la ausencia temporal de la señal, un *drop-out* de las cintas, acumulación de polvo en la superficie del CD, una interferencia producida por un pico de tensión, entre otros.

2) Los **errores aleatorios**, que provocan la pérdida de muestras aisladas. Suelen ser consecuencia de un ruido o de una calidad de señal insuficiente.

Para resolver o minimizar la acción negativa de estos errores, antes de grabar o transmitir la señal se intercalan una serie de bits de control en los datos de audio original que alteran el orden de las muestras iniciales. Un ejemplo de método de resolución de errores aleatorios sería el código Reed Solomon de intercalación cruzada (*cross interleave Reed Solomon code* o CIRC), que es el que se utiliza en los discos compactos.

El código Reed Solomon de intercalación cruzada (CIRC)

El CIRC hace sumas verticales, horizontales y diagonales de filas y columnas de las muestras cuantificadas y puede corregir más de una ráfaga de errores en una palabra.

A grandes rasgos, este código de corrección de errores funciona así: el sistema toma como bloque inicial seis periodos de muestreo. Cada periodo de muestreo tiene 16 bits que corresponden al canal izquierdo y 16 al derecho. Es decir, cada bloque de audio original está formado por doce palabras de 16 bits y un total de 192 bits ($32 \text{ bits} \times 6 \text{ periodos}$).

Cada grupo de 16 bits se separa en dos palabras de 8 bits. Por lo tanto, los 32 bits de cada periodo de muestreo pasan a estar expresados en cuatro palabras de 8 bits. El bloque de audio original, los seis periodos, pasa a ser veinticuatro palabras de 8 bits.

Estas veinticuatro palabras de 8 bits entran en el codificador CIRC, que genera ocho palabras de paridad y una palabra de principio de bloque, todas de 8 bits. Salen, por lo tanto, treinta y tres palabras de 8 bits ($24 + 8 + 1$).

A cada palabra de 8 bits se añaden 6 bits de intercalación para evitar ráfagas seguidas de 1 y 0. Esto se conoce como **modulación de bits 8 a 14** (*eight to fourteen modulation*, EFM). Ahora tenemos treinta y tres palabras de 14 bits.

El paso siguiente es intercalar un grupo de 3 bits de separación (*merging*) para unir cada palabra de 14 bits con la siguiente. Las treinta y tres palabras tienen 17 bits cada una. Suman un total de 561 bits.

El procesamiento acaba añadiendo 27 bits de sincronía al final del bloque. Por lo tanto, salen 588 bits.

En resumen, un bloque de audio original formado por 192 bits acaba teniendo 588 bits. O, dicho de otro modo, 192 bits representan la codificación del sonido original y 396 bits se encargan de la preservación y la seguridad.

Cuando se pierden muestras, estos datos redundantes de seguridad permiten interpolar las dos muestras adyacentes y obtener unos valores aproximados que llenen el vacío. Esta forma de disimular los errores con algoritmos matemáticos se conoce como cancelación. En un mensaje musical, la interpolación momentánea no provoca ninguna consecuencia perceptible. Otros sistemas más avanzados hacen un análisis en frecuencia del sonido y, cuando un valor se pierde, se insertan muestras que tengan las mismas características espectrales. Cuando se ha perdido mucha información y la cancelación no resuelve el problema, el sistema enmudece.

2.6. La señal digital electrónica

Después de multiplexar y aplicar los códigos de seguridad, los ceros y unos se transforman en pulsos eléctricos que se pueden almacenar en un soporte, enviar por una línea o transmitir radioeléctricamente.

Algunos códigos electrónicos

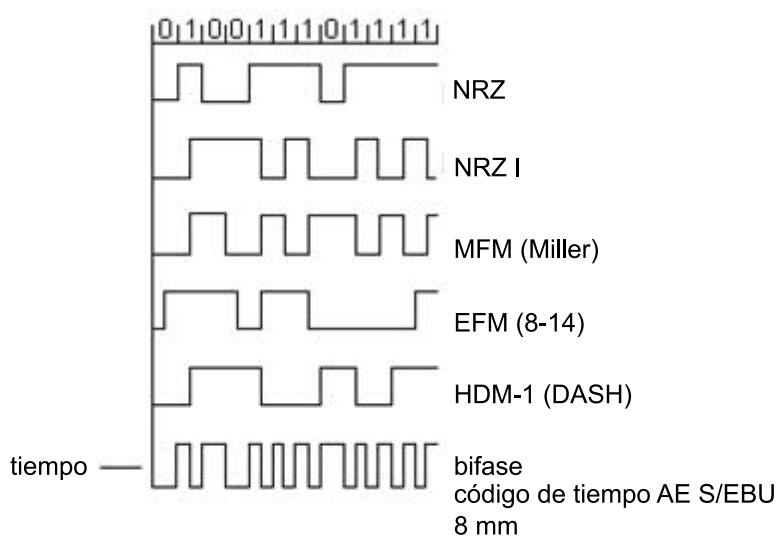
Se usan varios códigos electrónicos para representar los ceros y unos del código binario. Mencionaremos tres como ejemplo:

1) **Non return zero L**: da un nivel de voltaje diferente en cada bit. 1, voltaje positivo. 0, ausencia de voltaje.

2) **Non return zero polar**: da el mismo nivel de voltaje en cada bit, pero de signo contrario. 1, voltaje positivo; 0, voltaje negativo.

3) **Código RZ binario**: las señales de no retorno a 0 (NRZ) tienen el problema de que, cuando las series de uno o cero son muy largas, no se producen transiciones y esto dificulta la recuperación del sincronismo. Por eso, se han pensado unos códigos que hagan volver la señal a cero a medio intervalo. También hay dos variantes, polar y unipolar.

Otros códigos electrónicos son los EFM (*eight to fourteen modulation*), MFM (Miller) o HDM-1 (DASH).



El sistema electrónico digital tiene un umbral, de forma que si pasa la señal equivale a 1 y, si no pasa, a 0.



Finalmente, se aplican unas informaciones de sincronía y dirección a los impulsos electrónicos, equivalentes a 1 y 0, que permiten diferenciar el significado de los datos en cadenas de 0 y 1 consecutivos. Este proceso se conoce como **modulación**.

Una de las más empleadas es la **modulación por código de impulsos** o PCM⁵, creada en 1967. Otra modulación digital es la FSK⁶ o clave de desplazamiento de frecuencia.

⁽⁵⁾Sigla de *pulse code modulation*.

⁽⁶⁾Sigla de *frequency shift keying*.

Philips y Sony han desarrollado para los sistemas más avanzados, como el Super Audio CD, la tecnología alternativa Direct Stream Digital, que graba los pulsos de la señal digital bit a bit, aplicando una modulación delta-sigma y una frecuencia de muestreo sesenta y cuatro veces superior a los 44,1 kilohercios (64×44.100 hercios = 2.822 megahercios). Cada bit es una muestra.

3. Conversor digital/analógico

Un conversor digital/analógico restaura la señal digitalizada en el dominio analógico. El conversor digital/analógico da una señal continua a la salida que cambia de tensión según el valor de cada palabra binaria.

Una vez recuperados los impulsos electrónicos, se obtiene el 0 o 1 correspondiente de cada uno. Para recuperar la señal, solo hay que detectar el nivel; la forma de onda no es relevante. Así, aunque un impulso llegue deteriorado, si supera un cierto nivel da 1 y, si no lo supera, 0.

Lo que sí puede afectar negativamente al proceso son los cambios en la duración de los impulsos, las variaciones en el ancho. Este error en la base de tiempo se conoce con el nombre inglés de *jitter*. Si esta fluctuación temporal es elevada provoca clics o pequeñas interrupciones en la reproducción.

El problema se resuelve con la ayuda de un circuito de sincronía. Los datos se graban en una memoria RAM y después se leen a una cadencia fija controlada por un reloj de cristal de frecuencia extremadamente estable.

Cuando ya se han obtenido los bits, se les aplica el código de corrección de errores para cancelar las equivocaciones detectadas. Si la señal estaba multiplexada, se separa la información de cada canal. A continuación, cada palabra binaria entra en el conversor digital/analógico y, una vez descodificada, sale el nivel del voltaje de la señal de audio original de aquel instante.

El conversor digital/analógico tiene que ser lo más preciso posible. Por eso, algunos sistemas usan conversores de 20 o 24 bits, aunque las muestras procesadas sean de 16 bits. Así consiguen una resolución más buena.

También se aplica un sobremuestreo⁷ de la señal (*oversampling*) en el momento de la descodificación para tener más precisión en la reproducción de las frecuencias altas. Si no se hace el sobremuestreo, es necesario que la señal final pase por un filtro pasa bajo, muy preciso y caro –con frecuencia de corte de 20 kilohercios–, para rechazar las frecuencias de imagen que aparecen en torno a los 20 kilohercios durante el proceso.

⁽⁷⁾El sobremuestreo

El sobremuestreo consiste en descodificar las palabras binarias en ráfagas de bits más cortas, mediante el incremento de la frecuencia de exploración.

En un reproductor de disco compacto que hace un sobremuestreo de ocho veces, el conversor de digital a analógico no procesa 44.100 palabras de 16 bits por segundo, sino 352.800 ráfagas de 2 bits cada segundo ($44,1 \text{ kilohercios} \times 8 = 352,8 \text{ kilohercios}$).

Se pueden aplicar sobremuestreos de 4, 8, 16, 32, 64, 128 o 256 veces.

Hay dos tecnologías principales de conversores digital/analógico: la multibit y la de número bajo de bits (*low bit*).

Un **conversor multibit** procesa todos los ceros y unos de cada palabra binaria en paralelo. El circuito puede estar formado por una red de resistencias distribuidas en ramas paralelas, tantas como bits tenga la palabra binaria. Cada una da una intensidad ponderada según el peso del bit que le corresponde. El valor de la corriente de cada rama se duplica a medida que el peso del bit aumenta. La suma de todas estas intensidades da el nivel de la señal analógica. Estos conversores también se denominan $R/2R$ por la relación que tienen los valores de las resistencias que se usan. Otros sistemas multibit usan circuitos basados en condensadores o en amplificadores diferenciales. A la salida siempre se obtiene una tensión proporcional a los valores binarios de la entrada.

Sea cual sea la tecnología, todos los valores de los elementos de estos circuitos tienen que estar muy bien ajustados y los componentes deben tener una tolerancia eléctrica muy pequeña para garantizar una distorsión armónica muy baja y una respuesta en frecuencia lineal.

Cuando los procesadores informáticos empezaron a incrementar la velocidad de trabajo surgieron los **conversores de número bajo de bits**. Estos circuitos no tienen un ajuste tan delicado como los anteriores.

Los **conversores de número bajo de bits** consiguen resultados sonoros más buenos, sobre todo cuando el nivel de la señal es bajo.

El conversor Philips Bitstream

El conversor Philips Bitstream de 1 bit procesa los 16 bits de cada muestra de un CD uno a uno, en serie.

Configuraciones típicas y técnicas principales de los conversores de número bajo de bits

Para palabras de 16 bits, las configuraciones típicas que usan los conversores de número bajo de bits son de 1 bit con una frecuencia de trabajo de 11,3 megahercios (sobremuestreo de 256 veces) y de 4 bits a 2,8 megahercios (sobremuestreo de 64 veces).

Dentro de los conversores de número bajo de bits hay un par de técnicas principales, la Pulse Width Modulation, desarrollada por Matsushita, y la Pulse Density Modulation, de Philips.

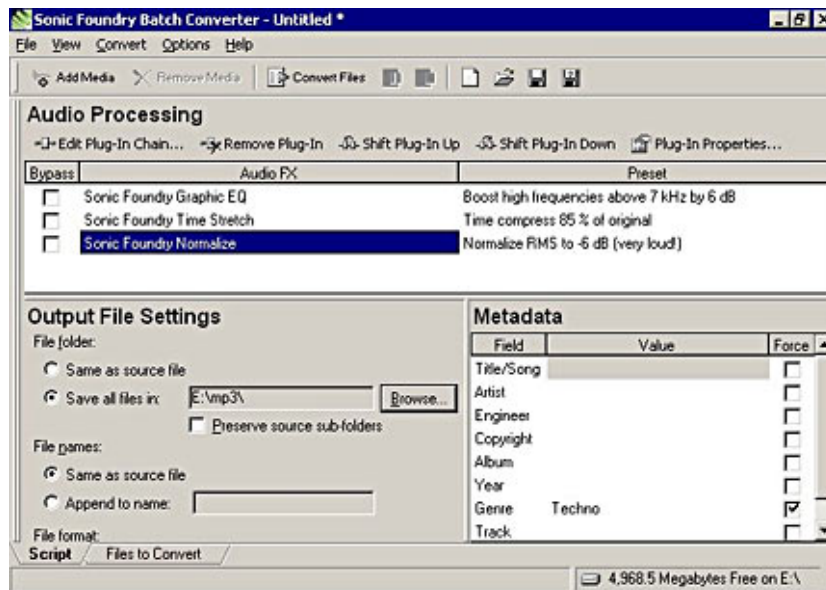
Aprovechando las elevadas velocidades de los procesadores informáticos desde mediados de los años noventa, la tecnología de los conversores de número bajo de bits ha evolucionado y ha dado el sistema Direct Stream Digital (DSD). El DSD es la base del formato Super Audio CD, de Sony y Philips, y procesa la señal digital con una frecuencia de conversión muy alta.

Para evitar las discontinuidades a la salida del conversor digital analógico, hay un circuito de retención y medición (*sample and hold*) que conserva el valor de la última muestra descodificada hasta que la palabra binaria siguiente haya entrado en el sistema.

Finalmente, la señal analógica pasa por un filtro pasa bajo que suaviza la forma de la onda.

4. Formatos de audio

Hay muchos formatos de audio digital. Para poder hacer el máximo número de transferencias o conversiones, las estaciones de audio digital tienen normalmente un área llamada *batch processing* o *batch converter*, que permite convertir formatos, frecuencias de muestreo o número de bits de cuantificación.



© Sonic Foundry, Inc.

Los principales formatos de audio son los que se presentan en la tabla siguiente:

wav (wave)	Creado por Microsoft e IBM como una evolución del formato RIFF 8 (<i>resource interchange file format</i>). Fue el primer formato de audio para PC. El formato wav/PCM (<i>pulse code modulator</i>) no está comprimido. Puede cuantificar a 8 o 16 bits. La frecuencia de muestreo puede llegar a los 44,1 kilohercios. Por lo tanto, tiene una calidad igual a la de un CD.
aiff (audio interchange file format)	Fichero de Apple para sus ordenadores Mac. Es una evolución de los ficheros IFF que usaban los Amiga. El fichero AIFF no comprime el sonido y trabaja con palabras de 8 bits. Tiene una calidad similar a la de un CD. También existe el formato AIFF-C, que está comprimido.
au (audio)	Fichero de Sun Microsystems que puede cuantificar a 8 o 16 bits. Es uno de los formatos más veteranos de los ficheros de sonido para Internet.

bwf (broadcast wave format)	En 1987, la Unión Europea de Radiodifusión definió el formato <i>bwf</i> para intercambios de sonido entre radios (EBU documento 3285). El <i>broadcast wave format</i> es un fichero de sonido basado en el formato <i>wave</i> que incluye una mínima información sobre el contenido (metadato). La versión 1 del <i>bwf</i> destina 64 bytes de los 254 reservados a datos a la información SMPTE UMID (<i>unique material identifier</i>).
Otros	Otras extensiones de ficheros de sonido que podemos encontrar son voc (Soundblaster), snd (de Sun), sf (de IRCAM), vox (dialógico), mat (<i>matlab</i>), SDD II (Sound Designer 2 de Digidesign) o raw (fichero <i>audio frame</i> en bruto sin ningún formato concreto).

4.1. Formatos de compresión de audio

Comprimir un sonido digital consiste en reducir el espacio que ocupa para no tener que procesar tantos bits. Así se consigue más capacidad de almacenamiento y una reducción del ancho de banda de la transmisión de audio.

El proceso de compresión consiste en hacer pasar la señal digitalizada por un compresor de datos (*encoder*) antes de la grabación o la transmisión y hacerlo pasar por un descompresor (*decoder*) antes de reproducirlo.

Cita

Tal como se explica en el temario de la asignatura de la UOC *Fundamentos de tecnología audiovisual*,

"La compresión está basada en la eliminación sucesiva de elementos redundantes que aparecen en la señal de audio y si se eliminan en el codificador es posible reconstruirlos en el decodificador".

"La redundancia estadística consiste en determinar cuántos valores de bits se repiten más en una secuencia de bits. La base de la eliminación de la redundancia estadística recae en el hecho de que hay palabras (agrupaciones de bits) que se repiten constantemente. Por lo tanto, la idea es enviar estas palabras que se repiten con menos bits."

Existen dos **sistemas de compresión**:

1) **Sin pérdidas de información (*lossy*)**, que reducen la ocupación entre un 30% y un 50% con métodos estadísticos

2) **Con pérdidas de información (*lossless*)**, que aplican criterios psicoacústicos y consiguen reducciones de entre un 40% y un 80%. En este caso, se tienen en cuenta las limitaciones del oído humano y se elimina toda la información que no es perceptible. El sistema ATRAC⁸ empleado por Sony en el minidisc o el sistema de reproducción multicanal SDDS son ejemplos.

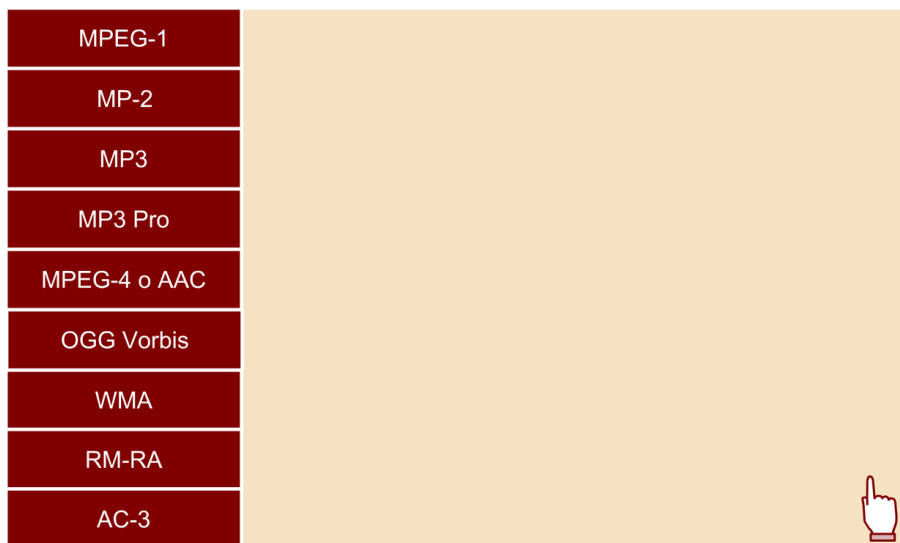
⁽⁸⁾ Acrónimo de *adaptive transform acoustic coding*.

Cuando una señal comprimida se recupera sin pérdidas, la señal obtenida es idéntica a la original. Por el contrario, si la compresión ha sido con pérdidas, el fichero descomprimido es, evidentemente, diferente al original. La compresión sin pérdidas solo es posible para aplicaciones que no sean en tiempo real.

La compresión de audio digital empieza en 1987 en los trabajos de investigación del Fraunhofer Institut Integrierte Schaltungen de Alemania. Aquellos algoritmos de compresión y descompresión fueron aceptados finalmente como estándar en 1992 por el MPEG⁹ y la ISO¹⁰ con el número ISO/IEC 11172. Nació la saga de sistemas MPEG.

⁽⁹⁾Sigla de Grupo de Expertos en Imágenes en Movimiento (en inglés, *Moving Picture Experts Group*).

⁽¹⁰⁾Sigla de Organización Internacional de Normalización (en inglés, *International Organization for Standardization*).



11

(11)

Calidad de audio	Canales	Bits por segundo	Relación de compresión
Peor que telefónica	mono	8 kbps	96:1
Telefónica	mono	16 kbps	48:1
Radio en AM (onda media)	mono	32 kbps	24:1
Similar a la FM	estéreo	64 kbps	24:1
FM de buena calidad	estéreo	96 kbps	16:1

Calidad de audio	Canales	Bits por segundo	Relación de compresión
Calidad CD	estéreo	Más de 128 kbps	Menos de 12:1

4.2. Formatos de distribución musical

Existen varios formatos de audio diseñados para la distribución y venta de música por Internet de manera segura. Algunos han tenido una vida muy corta –como el A2B de AT&T– y otros todavía se usan. Recogemos algunos ejemplos de un terreno en el que trabajan las discográficas para encontrar una nueva forma de vender los productos.

Liquid Audio (LQT)	Representa una alternativa al formato MP3 al garantizar más seguridad en la distribución de audio a las compañías discográficas. El reproductor de Liquid Audio se puede conseguir de forma gratuita en todas las páginas web que utilizan este sistema.
VQF	Desarrollado por Yamaha. Los archivos ocupan un 30% menos que los MP3. Se puede leer con el reproductor Winamp, añadiendo un conector. Solo permite acceder a las canciones si se conoce una clave.
SAF (Secure Audio Format)	En el verano del año 2000, BMG y Universal crearon con InterTrust el sistema Digibox, que garantizaba el pago y el respeto al copyright de las canciones o los discos enteros que el cliente quisiera comprar. Finalmente, el sistema se llamó SAF (<i>secure audio format</i>). Comprime el audio en formato MPEG-4 o AAC (<i>advanced audio code</i>).
RMF (<i>rich music format</i>)	Formato de Beatnik de descarga de audio de alta calidad.