

3. El audio digital en el ordenador personal

3.1 Introducción

Hemos visto en el capítulo anterior que para trabajar con sonido digital en un ordenador, son necesarios como mínimo dos conversores (A/D y D/A) que cumplan la función de digitalizar y reconstituir posteriormente la señal analógica. En los IBM PC compatibles esta posibilidad llegó hace unos años de la mano de las primeras tarjetas de sonido de tipo Sound Blaster, inicialmente con una calidad bastante lamentable, pero en otros ordenadores ya era posible desde años atrás. En esta obra nos centraremos básicamente en las posibilidades de los ordenadores PC compatibles actuales, dotados con una tarjeta de sonido de 16 bits, pero antes de entrar en materia repasaremos brevemente la historia del audio digital en los ordenadores.

3.2. Un poco de historia

3.2.1. Los inicios

Como se comenta en el capítulo anterior, Max Mathews consiguió sintetizar sonido digital en un ordenador ¡a finales de la década de los cincuenta! Aunque se tardó unos años más en disponer de la velocidad suficiente para muestrear sonido analógico, las posibilidades abiertas por Mathews iniciaron dos largas décadas de investigaciones en el terreno de la síntesis digital, llevadas a cabo inevitablemente en ordenadores mainframes y estaciones de trabajo Unix, en centros universitarios o laboratorios de importantes empresas informáticas. Volveremos a tratar este tema en el capítulo 9, “Síntesis y generación digital de sonido”.

3.2.2. Audio digital para las masas

El Commodore Amiga fue el primer ordenador doméstico que incorporó, en 1985, ciertas posibilidades de audio digital. El Amiga disponía de cuatro conversores D/A de 8 bits. Carecía de conversores A/D, por lo que si no se compraba un dispositivo muestreador adicional conectable al puerto paralelo, sólo funcionaba como reproductor. Estos primeros conversores D/A adicionales eran realmente muy económicos¹, pero tenían una limitación muy importante: por motivos de velocidad sólo permitían muestrear en memoria RAM, y no directamente al disco duro, por lo que la duración de los sonidos almacenables era muy

¹ Entre 5.000 y 15.000 ptas.

limitada. Aun con todas estas limitaciones, que lo hacían inviable para un trabajo profesional, el Amiga supuso para muchos la entrada al sonido digital².

Las primeras aplicaciones serias se dieron a finales de los ochenta en dos plataformas que ya de por sí, no eran nada baratas (Apple Macintosh y NeXT), mientras que el coste del hardware adicional superaba fácilmente el medio millón de pesetas. No eran productos “para las masas”, pero los estudios de grabación y algunos músicos profesionales se los pudieron ya permitir. Especialmente en el caso del Mac, el hardware y el software de Digidesign, que ofrecía cuatro canales de 16 bits, se convirtió (y sigue siendo hoy) en un estándar de la grabación y edición de audio digital en disco duro.

3.2.3. ... Y los PCs comenzaron a sonar

El bautizo sonoro del PC fue realmente tímido y tardío. En 1988, una compañía canadiense, Ad Lib, fabricó las primeras tarjetas de sonido. Soportaban tan solo cuatro voces con sonidos sintéticos y, aunque muchos juegos comenzaron a tener en cuenta el nuevo hardware, los tipos de sonidos ofrecidos por el pequeño sintetizador³ permitían poco más que generar una cierta musiquilla de fondo, con una calidad sonora bastante infame. Al no disponer de sonido digitalizado, la Ad Lib no favorecía la reproducción de efectos convincentes (golpes, motores, etc.) o voces (o gritos), que el mercado de los juegos estaba esperando.

Un año más tarde, Creative Labs, una empresa de Singapur que había trabajado en síntesis de voz, lanzó al mercado la primera Sound Blaster, que añadía a la compatibilidad con la Ad Lib unas pequeñas posibilidades de audio digital. Pese a que estas posibilidades eran limitadas (dos canales de 8 bits) el éxito fue tremendo y propulsó a la compañía al primer puesto del mercado multimedia internacional, lugar que sigue ocupando en la actualidad.

Aunque las tarjetas de hoy han cambiado bastante desde las Sound Blaster iniciales, muchas de las mejoras y variaciones están teniendo lugar en el terreno del MIDI, que trataremos en la segunda parte de este libro. En lo referente a audio digital, casi todas las tarjetas actuales ofrecen las mismas prestaciones desde aproximadamente 1994: dos canales de audio a 16 bits y hasta 44.100 Hz de frecuencia de muestreo, con posibilidad de grabación y reproducción a disco duro.

3.3. ¿Por qué la calidad CD no es siempre calidad CD?

Existe la creencia popular de que en el sonido digital no hay mejores ni peores, “como todo son ceros y unos, mientras no se confundan unos con otros ...”. Sin embargo, nada más lejos de la realidad. Los fabricantes de equipos multimedia se escudan detrás de los 16 bits y 44.100 Hz, para colocar la indiscriminada etiqueta “calidad CD”. Pero, ¿se ha preguntado alguna vez por qué existen actualmente lectores de CD-ROM por 5.000 ptas., mientras que se pueden adquirir lectores de CD Audio (que, por cierto, son siempre, como veremos más adelante, de *velocidad simple*) por 100.000 ptas.? La respuesta está en los conversores D/A

² Con estos sonidos almacenados en memoria, el Amiga podía funcionar como un *sampler* (véase apartado 9.6).

³ En el capítulo 9, “Síntesis y generación digital de sonido”, se trata a fondo el tema de los sintetizadores.

comentados en el anterior capítulo, y en otros componentes electrónicos analógicos, que son los responsables finales del sonido. Es cierto que el láser no suele confundir los ceros con los unos, pero para que estos enteros binarios lleguen a sonar, tienen que pasar por un complicado proceso de conversión, que se puede realizar con muy diversos criterios de calidad.

Un factor a tener en cuenta a la hora de evaluar la calidad de un dispositivo de audio digital, como una tarjeta de sonido, puede ser la relación señal/ruido comentada en el capítulo anterior. Aunque muchas tarjetas no incluyen estas especificaciones, es frecuente encontrarlas en los estudios comparativos realizados en revistas especializadas. Comentamos que el máximo teórico alcanzable con 16 bits de resolución es de 96 dB; una tarjeta con niveles inferiores a 80 dB debería ser descartada por demasiado “ruidosa”.

Sin embargo, muchas otras prestaciones importantes a la hora de evaluar actualmente una tarjeta de sonido, vienen dadas por sus capacidades MIDI. Por ello, en el capítulo 11, “El ordenador MIDI y la tarjeta de sonido”, trataremos con detalle la elección de este componente tan fundamental.

3.4. Tamaños en el audio digital

Cuando los ordenadores sólo manejaban texto, un megabyte (aproximadamente un millón de caracteres) parecía una cifra monstruosa. Hoy en día, con la llegada del multimedia, los sonidos y especialmente las imágenes digitales, han disparado estas cifras.

¿Cuánto ocupa un segundo de sonido digital estéreo de 16 bits y 44.100 Hz?

El cálculo es sencillo: en un sonido de 16 bits, cada muestra ocupa dos bytes (un byte son ocho bits), y si la frecuencia de muestreo es de 44.100 Hz, significa que cada segundo requiere de 44.100 muestras. Si el sonido es estéreo, utiliza dos canales, por lo que estas necesidades se ven duplicadas.

$2 \text{ bytes/muestra} \times 44.100 \text{ muestras/segundo} \times 2 \text{ (canales)} = 176.400 \text{ bytes/segundo} \text{ ó } 172,2 \text{ Kb/s.}$

Esta es forzosamente la velocidad de transferencia de los lectores de CD Audio, y es la que se designó en su día como *velocidad simple*. A partir de aquí, se tomó la costumbre de medir la velocidad de los lectores de CD-ROM en múltiplos de este valor (igual que la velocidad de los aviones reactores se mide a veces en múltiplos de la velocidad del sonido).

Realizando una multiplicación más, se observa que *un minuto de sonido digital estéreo de calidad, ocupa un valor muy cercano a los 10 Mb*. Sabemos por experiencia que los CD Audio raramente superan los 70 minutos. Esta duración límite aproximada viene dada por su capacidad, que suele ser de 720 Mb.

Cuando el sonido no es estéreo, estos tamaños descienden a la mitad, y lo mismo sucede si la resolución es de 8 en lugar de 16 bits, o la frecuencia de muestreo es de 22.050 Hz. Por ello, limitando la calidad al mínimo, el tamaño necesario para un minuto de sonido mono de 8 bits y 11.025 Hz, se reduce aproximadamente a 646 Kb. (estas condiciones suelen designarse como calidad telefónica).

3.4.1. ¿Qué calidad seleccionar?

No existe una respuesta unívoca a esta pregunta, ya que todo depende del uso que queramos dar a cada sonido. Sí que es posible indicar, a modo de guía, el orden de los pasos a seguir si necesitamos ahorrar memoria o espacio de disco duro. La reducción de calidad menos perceptible es utilizar 22.050 Hz en lugar de 44.100 Hz. Si el efecto estéreo no es importante, puede también limitar el sonido a un único canal, pero manteniendo los 16 bits de resolución. Con esto nos situamos a 42,05 Kb/segundo, y es el *mínimo* aconsejable para un sonido “musical”. Si lo que desea es grabar mensajes de voz, puede seguir reduciendo la frecuencia, o pasar directamente a los 8 bits. El que se indiquen valores de frecuencia de 44.100, 22.050 ó 11.025, no es casual, ya que éstos son los valores estándar que soportan la mayoría de las tarjetas (aunque algunas soporten también frecuencias intermedias).

Dado que hasta hace poco la mayoría de las tarjetas eran de 8 bits, muchos CD-ROMs interactivos utilizan todavía esta resolución, más por criterios de compatibilidad que de economía a ultranza.

3.5. Formatos de sonido digital

La información incluida en un fichero de sonido digital no es más que un *array* de bytes o de enteros de 16 bits, dependiendo de la resolución aplicada. Pero a pesar de la sencillez de su contenido, la lista de formatos existentes es inmensa, pues cada plataforma, y a veces cada programa dispone de una serie de formatos preferidos con ligeras variaciones entre uno y otro. Las diferencias pueden radicar en el tipo aplicado a las muestras (por ejemplo enteros con o sin signo), el tipo de compresión utilizado (si es que lo hubiera) y la forma en que estas muestras se ordenan. Para un sonido estéreo algunos formatos optan por guardar cada canal de forma independiente, mientras que otros eligen guardar muestras alternas de cada una de las pistas. Todos los formatos incluyen una cabecera (en la que se indica la resolución, la frecuencia de muestreo, el número de canales, etc.) cuyo tamaño, contenido y ordenación varía también de un formato a otro.

En Windows, el formato estándar es el WAV, aunque algunos fabricantes disponen también de formatos nativos (como es el caso del VOC de Creative Labs, utilizado en las primeras tarjetas Sound Blaster). Esta extensión es una abreviación de la palabra inglesa *wave*, que significa ola u onda; por ello este tipo de ficheros se conocen también como ficheros de onda. Si sólo trabaja con PC compatibles probablemente nunca tenga que preocuparse de otros formatos, salvo si desee importar ficheros procedentes de otras plataformas. En el entorno Macintosh, uno de los formatos más extendidos es el AIF, algo que deberán tener muy en cuenta los programadores de aplicaciones multimedia multiplataforma. En Internet se utiliza mucho el formato AU, propio de los sistemas Unix.

También es posible trabajar con ficheros sin cabecera (suelen venir con las extensiones PCM o RAW). Estos constituyen un caso particular, ya que a la hora de abrir un fichero de estas características, deberemos indicar nosotros la resolución, la frecuencia de muestreo y el número de canales. Si el fichero suena bien, habremos acertado. Si suena mal, habrá que volver a probar con parámetros diferentes.

Los programas de edición de audio que se comentan en el capítulo 5, ofrecen normalmente opciones para cargar y salvar formatos diferentes. En cualquier caso conviene tener en cuenta que dos ficheros no comprimidos, y con la misma resolución y frecuencia de muestreo, contendrán la misma información (aunque organizada de diversas formas) y sonarán exactamente igual, independientemente del formato. En la tabla 3.1 se enumeran algunos de estos formatos más frecuentes.

| Extensión | Plataforma - aplicaciones |
|-----------|--------------------------------------------------|
| AU | Next/Sun - Internet |
| AIF | Macintosh - Multimedia (<i>Director</i> , etc.) |
| IFF | Amiga |
| PCM | Cualquiera |
| RAW | Cualquiera |
| WAV | PC |

Tabla 3.1. Algunos formatos de ficheros de audio digital

3.6. Compresión de ficheros de audio

Dado el tamaño que pueden adquirir los ficheros de audio digital es lógico que se haya buscado formas de compresión que permitan reducir esta cantidad de información. Al igual que ocurre con la imagen, existen técnicas sin pérdida y técnicas con pérdida. Un factor importante en los sistemas de compresión de audio es que deben ser capaces de comprimir y descomprimir en tiempo real (para comprimir y descomprimir en diferido existen ya multitud de sistemas de uso general como el ZIP o el ARJ, que también pueden ser aplicados a ficheros de audio).

Las matemáticas involucradas en cualquier sistema de compresión son demasiado complejas para ser tratadas con rigor en esta obra, por lo que nos limitaremos a esbozar ciertas ideas. Muchos de estos sistemas son independientes de los formatos citados en el apartado anterior, de forma que un mismo formato de fichero puede soportar varios métodos de compresión.

- Probablemente le suenen las siglas ADPCM. Corresponden a *Adaptive Delta Pulse Code Modulation*, una forma de compresión de la que existen múltiples variantes, en la que se reduce el número de bits de la señal (por ejemplo a 4 bits) tratando únicamente las diferencia entre una muestra y la siguiente. Es fácil intuir que esto provoca una inevitable degradación de la señal ya que el sistema es incapaz de representar saltos grandes de amplitud entre dos muestras.
- Las compresiones μ -law y A-law, utilizan un sistema de compresión no lineal que permite utilizar una resolución de 8 bits, pero ofreciendo una calidad sonora y un rango dinámico próximos a los obtenidos con 14 bits. Son por lo general más rápidos que los métodos basados en ADPCM.
- Otros métodos como los utilizados por los grabadores digitales en formato DCC o MiniDisc, son variantes del sistema ADPCM, que eliminan de la señal original ciertos componentes que teóricamente no son audibles por quedar enmascarados (es como si en

una imagen 3D, eliminásemos toda la información de los objetos tapados por planos más próximos al observador).

Algunos de estos sistemas se ven favorecidos por la presencia de hardware especializado como pueda ser la inclusión de un DSP (procesador digital de señal) en la tarjeta de sonido, mientras que otros funcionan perfectamente por software. Si abre el icono de multimedia en el panel de control de Windows 95, y selecciona la forma de visualización **Avanzado**, podrá ver en el apartado de **Codecs⁴ de compresión de audio**, todos los compresores instalados en su sistema, tal como se observa en figura 3.1.

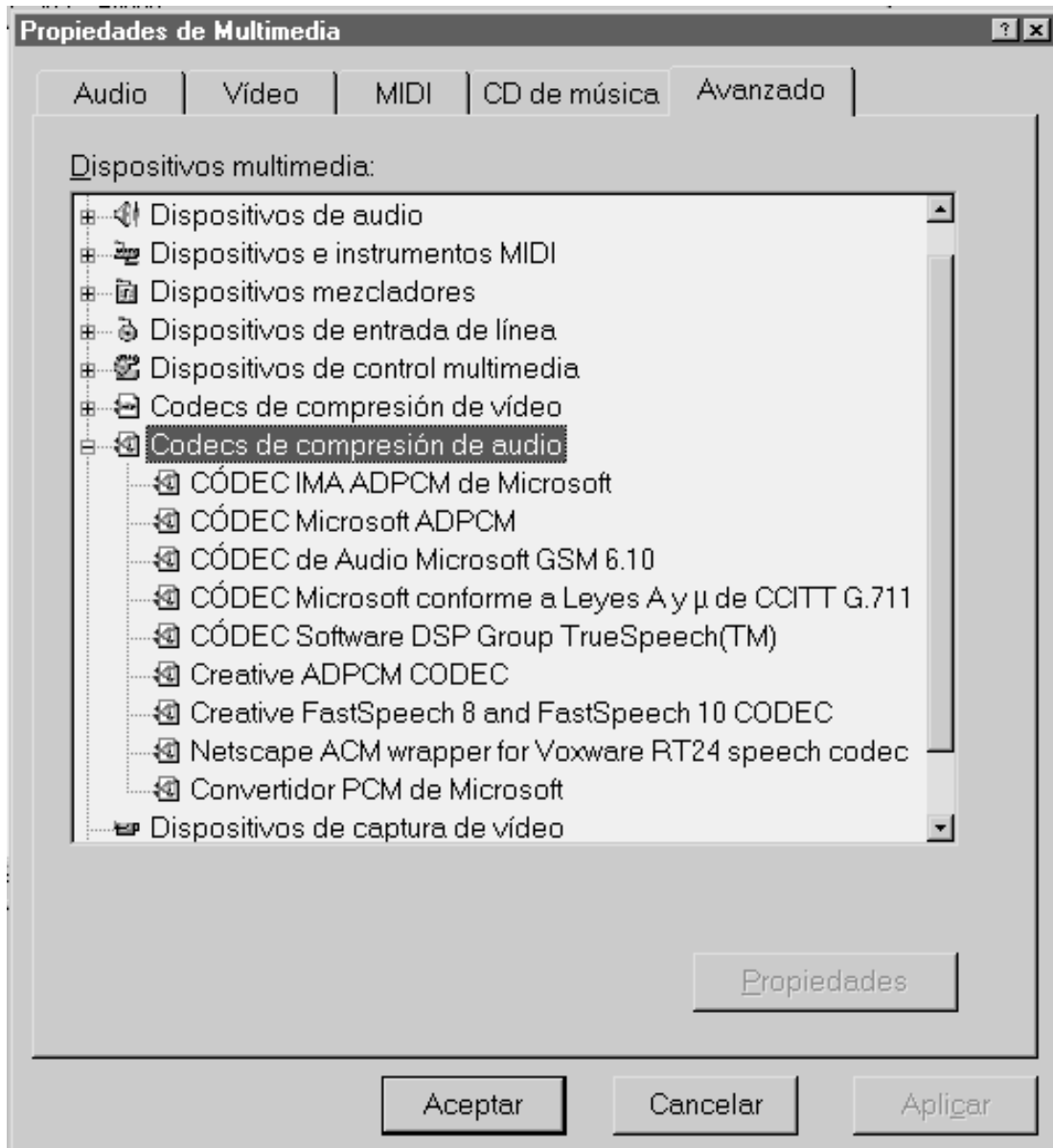


Figura 3.1. Comprobación de los *Codecs* de compresión instalados en el sistema

La compresión puede plantear problemas de compatibilidad entre diferentes ordenadores, por lo que sólo debería utilizarse en aquellos casos en los que el ahorro de espacio sea un imperativo. En la medida de lo posible, tampoco debería comprimir un fichero hasta que esté

⁴ El término *codec* proviene de la contracción de las dos palabras *codificador* y *descodificador*.

seguro de que no va a manipularlo más, pues aunque la pérdida pueda ser poco perceptible, siempre es preferible trabajar sobre la señal original.

3.7. Obtención de sonidos

A grandes trazos, existen dos formas de obtener un fichero de sonido digital: sintetizándolo desde el propio ordenador o digitalizando el sonido de una fuente externa. El primer método fue también el que, por menores requerimientos de potencia (la máquina no necesita adquirir la información en tiempo real) se implementó primero, hace ya cerca de cuarenta años. De él hablaremos en el capítulo 9, "Síntesis y generación digital de sonido". De momento, en el próximo capítulo "Audio digital en Windows 95", aprenderemos a configurar correctamente el sistema, y a realizar grabaciones digitales de diferentes fuentes externas, a partir de las posibilidades que nos brinda Windows 95.