



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 315 992**

51 Int. Cl.:
G10L 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06076046 .9**

96 Fecha de presentación : **30.04.2004**

97 Número de publicación de la solicitud: **1742203**

97 Fecha de publicación de la solicitud: **10.01.2007**

54 Título: **Control de nivel de audio para audio comprimido.**

30 Prioridad: **30.04.2003 US 426664**

45 Fecha de publicación de la mención BOPI:
01.04.2009

45 Fecha de la publicación del folleto de la patente:
01.04.2009

73 Titular/es: **The Directv Group, Inc.
2230 E. Imperial Highway
El Segundo, California 90245, US**

72 Inventor/es: **Michener, James A.**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 315 992 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Control de nivel de audio para audio comprimido.

5 Antecedentes del invento**1. Campo del invento**

El presente invento se refiere al control del nivel de audio para datos comprimidos.

10

2. Descripción de la técnica relacionada

15 La televisión digital, tal como la proporcionada por DIRECTV[®], la cesionaria del presente invento, se transmite típicamente como una corriente de datos digitales codificados utilizando la norma MPEG (Grupo de expertos en películas cinematográficas) promulgada por la ISO (Organización internacional para la estandarización). La MPEG proporciona una manera eficaz de representar video y audio en forma de una corriente de bits comprimidos.

20 La norma MPEG-1 se describe en un documento titulado "Codificación de imágenes en movimiento y audio asociado para medios digitales de almacenamiento hasta a, aproximadamente, 1,5 Mbit/s", ISO/IEC 11172 (1993). La norma MPEG-1 también se describe en un documento titulado "ISO-MPEG-1 audio: una norma genérica para codificar audio digital del alta calidad" de BRANDENBURG K y otros, JOURNAL OF THE AUDIO ENGINEERING SOCIETY, Octubre de 1994, paginas 780-792, XP000978167. La norma MPEG-2 se describe en un documento titulado "Codificación genérica de imágenes en movimiento e información de audio asociada", ISO/IEC-13818 (1998).

25 Aún cuando se trata de una emisora por satélite, DIRECTV[®] ofrece a sus abonados programación local, es decir, canales de televisión local, que requieren que cada uno de los canales de televisión en una ciudad se codifiquen en MPEG y se multiplexen estadísticamente en una instalación de recogida, antes de ser transportados, mediante una portadora común, a un centro emisor para envío por un enlace ascendente hacia satélites operados por DIRECTV[®]. Pueden cerrarse acuerdos con otras emisoras por satélite y operadores de cable para compartir estas instalaciones de recogida con el fin de reducir costes.

30 Además, los proveedores de programas, tales como Disney[®], Viacom[®], HBO[®], Showtime[®], Starz[®], ESPN[®], etc., proporcionan con frecuencia a DIRECTV[®] una corriente de datos MPEG previamente codificados y multiplexados estadísticamente. Estos proveedores de programas pueden pedir que la corriente de datos MPEG sea hecha pasar directamente a abonados de DIRECTV[®] sin descodificación y nueva codificación.

35 Sin embargo, pueden surgir problemas al utilizar estas distintas corrientes de datos MPEG debido a que los diferentes emisores por satélite, operadores de cable y proveedores de programas, pueden utilizar diferentes normas que den como resultado distintos niveles de audio. Por ejemplo, DIRECTV[®] sigue la recomendación de la SMPTE (Sociedad de ingenieros de películas cinematográficas y televisión), de que un nivel de referencia de 0 dB se encuentra a -20 dB de la escala digital completa, mientras que otras emisoras por satélite, operadores de cable o proveedores de programas pueden funcionar con un nivel de referencia de 0 dB que está a -17 dB de la escala digital completa.

40 Si estas diferentes corrientes de datos MPEG utilizan una o más normas diferentes, entonces los canales de emisión resultantes aparecerán demasiado fuertes o demasiado suaves, en comparación con otros canales. Así, existe la necesidad de cambiar los niveles de audio de una corriente de datos de audio MPEG.

45 Existen aplicaciones adicionales en las que es necesario disponer de la capacidad de cambiar los niveles de audio de una corriente de datos MPEG. Por ejemplo, la producción de televisión trabaja, generalmente, con un amplio margen dinámico, ofreciendo la posibilidad de que un programador creativo "suba" el audio durante escenas de clímax. Asimismo, la música clásica trabaja, con frecuencia, con un amplio margen dinámico.

50 Por otra parte, la música más popular tiene fuertemente limitado su margen dinámico. Esta limitación del margen dinámico se debe a muchas razones:

55

- 1) El artista desea que la música suene alta.
- 2) Las estaciones de radio creen, con frecuencia, que mantener un silencio es dejar de emitir.
- 60 3) En un ambiente con un elevado nivel de ruido, tal como en un automóvil, un estadio u otro recinto público, es necesario disponer de un estrecho margen dinámico para hacerse oír sobre el ruido.
- 4) La tecnología de grabación imparte un elevado nivel de ruido, por ejemplo, en las cintas de casete y un margen dinámico limitado enmascara el ruido.
- 65 5) La tecnología de reproducción tiene un margen dinámico limitado, por ejemplo, los dispositivos de audición personales activados por baterías.

ES 2 315 992 T3

En lo que respecta a los reproductores MPEG, el nivel de referencia de 0 dB para muchos de estos dispositivos se encuentra a -10 dB de una escala digital completa. En consecuencia, si una corriente de datos de audio MPEG utiliza un nivel de referencia de 0 dB a -20 dB de la escala digital completa, entonces el control de volumen del dispositivo tendría que ser incrementado en 10 dB para compensar. Sin embargo, en muchos de estos dispositivos existe un margen de ganancia limitado, ya que no soportan un audio con un gran margen dinámico. Entonces, una solución mejor es cambiar los niveles de audio de la corriente de datos de audio MPEG.

En la técnica anterior, un método de alterar los niveles de audio comprendería (1) descodificar (descomprimir) la corriente de datos de audio MPEG, (2) ajustar la ganancia, y (3) codificar (comprimir de nuevo) la corriente de datos de audio MPEG. Este método es ventajoso porque codificadores y descodificadores comercialmente disponibles pueden comprarse a precio relativamente bajo. Sin embargo, este método tiene muchos inconvenientes, incluyendo la inyección de un retardo de tiempo considerable, al menos 48 milisegundos (ms), así como un aumento del ruido y de la distorsión provocado por, todavía, otra nueva cuantización del audio.

En consecuencia, existe la necesidad de poder cambiar los niveles de audio de corrientes de datos de audio MPEG sin descomprimir los datos de audio dentro de las corrientes de datos de audio MPEG, alterar los niveles de ganancia de los datos de audio y, luego, volver a comprimir los datos de audio dentro de las corrientes de datos de audio MPEG.

Compendio del invento

Un objetivo de la invención es proporcionar un método de detección de nivel de audio para audio comprimido y un aparato para ello, tal como se reivindica en las reivindicaciones adjuntas.

Breve descripción de los dibujos

Nos referiremos ahora a los dibujos, en los que números de referencia similares representan partes correspondientes en todos ellos:

la Fig. 1 es un diagrama de bloques que ilustra un entorno ilustrativo utilizado para llevar a la práctica la realización preferida del invento;

la Fig. 2 es un diagrama de bloques que ilustra la estructura de una corriente de datos de audio MPEG; y

la Fig. 3 es una gráfica de flujo que ilustra la lógica ejecutada por un proceso de alteración de ganancia para cambiar los factores de escala sin alterar los datos de audio comprimidos en sub-bandas, con el fin de proporcionar un control del nivel de audio.

Descripción detallada de las realizaciones preferidas

En la descripción siguiente, se hace referencia a los dibujos adjuntos, que forman parte de la misma y que muestran, a modo de ilustración, varias realizaciones del presente invento. Ha de entenderse que pueden utilizarse otras realizaciones y que pueden llevarse a cabo cambios estructurales sin apartarse del alcance del presente invento.

Visión general

El presente invento se refiere al control del nivel de audio para audio comprimido. Específicamente, el presente invento se refiere a la extracción de factores de escala para el audio comprimido desde una corriente de datos de audio MPEG, la alteración de los factores de escala extraídos sin descomprimir el audio comprimido con el fin de proporcionar un control del nivel de audio, y la actualización de la corriente de datos de audio MPEG con los factores de escala alterados. Todos los factores de escala de la corriente de datos de audio MPEG son alterados basándose en un parámetro que identifica cómo han de alterarse los niveles de ganancia de la corriente de datos MPEG.

En consecuencia, si una corriente de datos de audio es demasiado fuerte o demasiado suave, el nivel de audio puede regularse en la forma deseada con el fin de mantener niveles de escucha uniformes. Esto supone una mejora con respecto a las técnicas anteriores que descomprimen los datos de audio, alteran los niveles de ganancia de los datos de audio y, luego, vuelven a comprimir los datos de audio, en las que el ciclo de descompresión y nueva compresión provoca el deterioro de la calidad de la señal y retarda el audio.

Entorno ilustrativo

La Fig. 1 es un diagrama de bloques que muestra un entorno ilustrativo utilizado para llevar a la práctica la realización preferida del invento. En el entorno ilustrativo, un procesador 100 puede incluir, entre otros, lógica, memoria y varios periféricos diferentes. De preferencia, el procesador 100 ejecuta un proceso 102 de alteración de ganancia, que realiza un cambio del nivel de audio, así como una detección del nivel de audio, directamente sobre una corriente de datos de audio MPEG, sin descomprimir y, luego, volver a comprimir los datos de audio dentro de la corriente de

ES 2 315 992 T3

datos de audio MPEG. Específicamente, el proceso 102 de alteración de ganancia acepta como entrada una corriente 104 de datos de audio MPEG, altera los factores de escala sub-banda encontrados en la corriente 104 de datos de audio MPEG, actualiza la corriente 104 de datos de audio MPEG con los factores de escala de sub-banda alterados y, luego, emite como salida la corriente 106 de datos de audio MPEG actualizada.

5
10
15
20
25
30
35
40
45
50

Generalmente, el proceso 102 de alteración de ganancia comprende lógica, instrucciones y/o datos que están incorporados en un dispositivo, medio, portador o señal, por ejemplo el propio procesador 100, o que pueden ser recuperados de él, una memoria, dispositivo de almacenamiento de datos o dispositivo remoto acoplado al procesador 100, etc. Además, esta lógica, instrucciones y/o datos, cuando son utilizadas, ejecutadas y/o interpretadas por el procesador 100, hacen que éste lleve a cabo los pasos necesarios para llevar a la práctica y/o usar el presente invento. En consecuencia, el presente invento puede ser incorporado como un método, aparato o artículo de fabricación que emplee software, firmware (soporte lógico inalterable) o cualquier combinación de ellos. Los expertos en la técnica reconocerán que en esta configuración pueden realizarse muchas modificaciones sin salirse por ello del alcance del presente invento.

15 *Corriente de datos de audio MPEG*

La Fig. 2 es un diagrama de bloques que ilustra la estructura de una corriente 200 de datos de audio MPEG. Capas I, II y III dentro de la corriente 200 de datos de audio MPEG se muestran como tramas separadas 202, 204 y 206.

20
25
30
35
40
45
50

Cada cuadro 202, 204 y 206 incluye una cabecera 206 que va seguida por una comprobación de redundancia cíclica opcional (CRC) 210 de 16 bits de longitud. La cabecera 208 es de 32 bits e incluye la siguiente información:

- vocablo de sincronismo - 12 bits (todos 1)
- vocablo de sistema - 20 bits
 - Identificador de versión - 1 bit
 - Capa - 2 bits
 - Protección contra errores - 1 bit
 - Índice de tasa binaria - 4 bits
 - Índice de tasa de frecuencia de muestreo - 2 bits
 - Relleno - 1 bit
 - Privado - 1 bit
 - Modo - 2 bits
 - Extensión de modo - 2 bits
 - Derechos de copia - 1 bit
 - Original o copia - 1 bit
 - Énfasis - 2 bits

La CRC 210, si está presente, es utilizada para detectar errores.

55
60

En el cuadro 202 de la capa I, la CRC 210 va seguida por una asignación de bit 212 (128-256 bits de longitud), factores de escala 214 (0-384 bits de longitud), muestras 216 (384 bits de longitud) y datos auxiliares 218. En el cuadro 204 de la capa II, la CRC 210 va seguida por una asignación de bit 212 (26-188 bits de longitud), información de selección de factor de escala (SCFSI) 220 (0-60 bits de longitud), factores de escala 214 (0-1080 bits de longitud), muestras 216 (1152 bits de longitud) y datos auxiliares 218. En el cuadro 206 de la capa III, la CRC 210 va seguida por información lateral 222 (136-256 bits de longitud) y un depósito de bits 224.

65

La asignación de bit 212 determina el número de bits por muestra para la capa I o el número de niveles de cuantización para la capa II. Específicamente, la asignación de bits 212 especifica el número de bits asignado para cuantización de cada sub-banda. Estas asignaciones se realizan de manera adaptable, de acuerdo con el contenido de información de la señal de audio, por lo que la asignación de bits 212 varía en cada cuadro 202, 204. Las muestras 216 pueden codificarse con bits cero (es decir, no hay datos presentes), o con entre dos y quince bits por muestra.

Los factores de escala 214 se codifican para indicar sesenta y tres posibles valores que se codifican como diseños de índice de seis bits, desde "000000" (0), que designa el factor de escala máximo, hasta "111111" (62) que designa

ES 2 315 992 T3

el factor de escala mínimo. Cada sub-banda de las muestras 216 tiene un factor de escala 214 asociado que define el nivel al que se recombina cada sub-banda durante la descodificación.

5 Las muestras 216 comprenden datos de audio comprimidos para cada una de treinta y dos sub-bandas. Un cuadro 202 de la capa I comprende doce muestras por sub-banda. Un cuadro 204 de la capa II comprende treinta y seis muestras por sub-banda.

10 En la capa II 204, las muestras 216 de cada cuadro se dividen en tres partes, cada una de las cuales comprende doce muestras por sub-banda. Para cada sub-banda, la SCFSI 220 indica si las tres partes tienen factores de escala 214 separados o si las tres partes tienen el mismo factor de escala 214, o si dos partes (las dos primeras o las dos últimas) tienen un factor de escala 214 y la otra parte tiene otro factor de escala 214.

15 Durante la descompresión, se proporcionan las muestras 216 a un cuantizador inverso, que selecciona valores predeterminados de acuerdo con la asignación de bits 212 y lleva a cabo una operación de descuantización, en la que los valores descuantizados son multiplicados, luego, por los factores de escala 214 para obtener valores no normalizados. Así, si se cambian todos los factores de escala 214 de sub-banda, se alterará el nivel de audio. Además, estos cambios de los factores de escala 214 pueden llevarse a cabo sin alteración de los datos de audio comprimidos de las sub-bandas.

20 *Lógica del proceso de alteración de ganancia*

La Fig. 3 es una gráfica de flujo que ilustra la lógica seguida por el proceso 102 de alteración de ganancia para cambiar los factores de escala 214 sin alterar los datos de audio comprimidos de las sub-bandas, de acuerdo con una realización preferida del presente invento. A este respecto, el proceso 102 de alteración de ganancia es un filtro en el que entra la corriente 104 de datos de audio MPEG, se alteran los factores de escala 214 y se actualiza la corriente 106 de datos de audio MPEG de salida con los factores de escala 214 alterados (pero que por lo demás no sufre más cambios con respecto a la corriente 104 de datos de audio MPEG). En la realización preferida, el proceso 102 de alteración de ganancia solamente incurre en una latencia de 2 bytes en su tratamiento, lo que provoca un retardo mínimo.

30 El bloque 300 representa el proceso 102 de alteración de ganancia aceptando un byte cada vez de la corriente 104 de datos de audio MPEG de entrada, así como un parámetro que identifica cómo han de alterarse los niveles de ganancia de la corriente 104 de datos de audio MPEG de entrada.

35 El bloque 302 representa la lógica de una sentencia CASE que es activada por un valor de estado corriente, en la que el control se transfiere a los bloques 304-322, dependiendo del valor de estado corriente. Tras haberse ejecutado la lógica de los bloques 304-322 para el estado corriente, el control se transfiere al bloque 324, que emite como salida un número de bytes indicado por los bloques 304-322 a la corriente 106 de audio MPEG de salida. Después de eso, el control retorna al bloque 300 para tratar el siguiente byte de entrada.

40 El bloque 304 representa un estado de 0. En este estado, el proceso 102 de alteración de ganancia espera hasta recibir el primer byte del vocablo de sincronismo procedente de la cabecera 208 de la corriente 104 de datos de audio MPEG de entrada. Específicamente, si el byte de entrada es igual a 0xff, entonces se incrementa el estado; de otro modo, no ocurre nada. Después, el control es transferido al bloque 324, que emite como salida el byte de entrada sin cambios.

45 El bloque 306 representa un estado de 1. En este estado, el proceso 102 de alteración de ganancia examina el byte de entrada para determinar si es el segundo byte a continuación del primero del vocablo de sincronismo procedente de la cabecera 208 de la corriente 104 de datos de audio MPEG de entrada, cuyo segundo byte incluye los 4 bits menos significativos del vocablo de sincronismo de 12 bits de la cabecera 208 y los 4 bits más significativos del vocablo de sistema de 20 bits de la cabecera 208. Si no es así, entonces el estado es devuelto a 0 y se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios. De otro modo se extraen los bits de protección contra errores y de capa de entre los 4 bits más significativos del vocablo de sistema de 20 bits de la cabecera 208 de la corriente 104 de datos de audio MPEG de entrada. Si la protección contra errores es 1 (activa) o si la capa no es 2 (capa II MPEG), entonces el estado es devuelto a 0 y se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios. (Obsérvese que esta realización solamente soporta audio de capa II MPEG sin protección). De otro modo, se incrementa el estado y se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios.

60 El bloque 308 representa un estado de 2. En este estado, el proceso 102 de alteración de ganancia extrae el índice de frecuencia binaria y el índice de tasa de frecuencia de muestreo de entre 8 bits adicionales del vocablo del sistema de 20 bits de la cabecera 208 de la corriente 104 de datos de audio MPEG de entrada. El índice de tasa binaria, junto con la capa (2) previamente extraída, se emplean como índice en una tabla de tasa binaria, que determina una tasa binaria. El índice de tasa de frecuencia de muestreo se utiliza como índice en una tabla de tasa de frecuencia de muestreo que determina una tasa de frecuencia de muestreo. Si la tasa de frecuencia de muestreo no es válida, entonces el estado es devuelto a 0; de otro modo, se incrementa el estado. Se transfiere entonces el control al bloque 324, que emite como salida el byte de entrada sin cambios.

ES 2 315 992 T3

El bloque 310 representa un estado de 3. En este estado, el proceso 102 de alteración de ganancia extrae el modo y la extensión de modo de los 8 bits finales del vocablo de sistema de 20 bits de la cabecera 208 de la corriente 104 de datos de audio MPEG de entrada. Con el modo y la extensión de modo, así como una tasa de frecuencia de muestreo obtenida a partir del estado 2, se determinan varias sub-bandas y varios canales para cada sub-banda. Se incrementa el estado y, luego, se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios.

El bloque 312 representa un estado de 4. En este estado, el proceso 102 de alteración de ganancia recoge el primer byte de la CRC 210 de la corriente 104 de datos de audio MPEG de entrada. Se incrementa el estado y, luego, se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios.

El bloque 314 representa un estado de 5. En este estado, el proceso 102 de alteración de ganancia recoge el segundo byte de la CRC 210 de la corriente 104 de datos de audio MPEG de entrada. Se incrementa el estado y, luego, se transfiere el control al bloque 324, que emite como salida el byte de entrada sin cambios. Obsérvese que los estados 4 y 5 recogerían la CRC 210 para recálculo posterior después de haber sido alterados los factores de escala 214. Sin embargo, en esta exposición se omite la descripción de la CRC 210.

El bloque 316 representa un estado de 6. En este estado, el proceso 102 de alteración de ganancia extrae la asignación de bit 210 de la corriente 104 de datos de audio MPEG de entrada. El número de bytes de entrada recibidos mientras se está en este estado es determinado por el número de sub-bandas y el número de modos. En consecuencia, el proceso 102 de alteración de ganancia se mantiene en este estado hasta que se ha recibido la totalidad de la asignación de bit 210. Hasta que eso ocurre, el estado no cambia y, luego, el control es transferido al bloque 324, que emite como salida el byte de entrada sin cambios. Una vez recibida la totalidad de la asignación de bit 210, se incrementa el estado y se transfiere, entonces, el control al bloque 324, que también emite como salida el byte de entrada sin cambios.

El bloque 318 representa un estado de 7. En este estado, el proceso 102 de alteración de ganancia extrae la SCFSI 220 de la corriente 104 de datos de audio MPEG de entrada. El tamaño del campo SCFSI 220 se basa en el número de sub-bandas y en la asignación de bit 210. En consecuencia, el proceso 102 de alteración de ganancia se mantiene en este estado hasta que se ha recibido la totalidad de la SCFSI. Hasta que eso ocurra, el estado no cambia y, luego, el control se transfiere al bloque 324, que emite como salida el byte de entrada sin cambios. Una vez recibida la totalidad de la SCFSI se incrementa el estado y, entonces, se transfiere el control al bloque 324, que también emite como salida el byte de entrada sin cambios.

El bloque 320 representa un estado de 8. En este estado, el proceso 102 de alteración de ganancia extrae los factores de escala 214 para cada sub-banda de la corriente 104 de datos de audio MPEG de entrada, cuyos factores de escala 214 comprenden multiplicadores para sub-bandas de los datos de audio. Una vez que se ha extraído un factor de escala 214, se le altera, por ejemplo se le incrementa o se le decrementa, de acuerdo con el parámetro que identifica cómo han de alterarse los niveles de ganancia de la corriente 104 de datos de audio MPEG de entrada.

Cada factor de escala 214 ocupa seis bits, que no están en alineación en un byte. En consecuencia, para alterar los factores de escala 214 hay momentos cuando los resultados de un byte de entrada previo tienen que ser retenidos para un byte de entrada adicional, antes de que puedan ser alterados y, luego, emitidos como salida. Mientras se están extrayendo los factores de escala 214, el estado se mantiene sin cambios y, luego, el control es transferido al bloque 324, que emite como salida el número de bytes para los factores de escala 214 alterados (0, 1 o 2) a medida que están disponibles.

Los factores de escala 214 son enteros que van de 0 a 63 y se utilizan como multiplicadores para la salida de sub-banda. Los factores de escala 214 alterados son limitados y no se inician de nuevo. En cambio, los factores de escala 214 alterados están limitados a 0 o 63, no disminuyendo los factores de escala 214 alterados por debajo de un mínimo (0) ni aumentando por encima de un máximo (63).

El que los factores de escala 214 alterados estén limitados mientras disminuye la ganancia quiere decir que un error ocurriría en un nivel de amplitud de -140 dB, que está muy por debajo del umbral de percepción auditiva. Por otra parte, el que los factores de escala 214 alterados estén limitados mientras crece la ganancia, quiere decir que todas las otras sub-bandas verán aumentada su amplitud, mientras esta sub-banda no puede aumentar tanto. Sin embargo, este efecto es, con frecuencia, muy notable, aunque no es probable que ocurra, ya que haría necesario incrementar el volumen hasta un nivel excesivamente alto, por ejemplo de unos 20 dB por encima del nivel medio.

Como se ha hecho notar en lo que antecede, el proceso 102 de alteración de ganancia continúa en este estado hasta que se han alterado todos los factores de escala 214, en cuyo momento se incrementa el estado y se transfiere el control, entonces, al bloque 324 que emite como salida el número de bytes para los últimos factores de escala 214 alterados restantes (1 o 2).

El bloque 322 representa un estado de 9. En este estado, el proceso 102 de alteración de ganancia no realiza función alguna. En consecuencia, el estado se mantiene sin cambios y el control se transfiere, entonces, al bloque 324, que emite como salida el byte de entrada sin cambios. El proceso 102 de alteración de ganancia continúa en este estado hasta que es repuesto externamente. De preferencia, el proceso 102 de alteración de ganancia es repuesto externamente basándose en el número de bytes de datos y leyendo la tasa binaria y la tasa de frecuencia de muestreo a partir de la cabecera MPEG.

ES 2 315 992 T3

Detección de nivel

Además de alterar el nivel de audio de la corriente de datos de audio MPEG, el presente invento también puede llevar a cabo una detección de nivel para el audio comprimido, cuya detección de nivel determina si todavía está presente audio. Esto ocurre porque los factores de escala 214 de la corriente de datos de audio MPEG representan un valor de pico del nivel de sub-banda por encima de los 24 ms de cada paquete de la corriente de datos de audio MPEG.

La detección de nivel para el audio comprimido supone: (1) obtener la raíz cuadrada de la suma de los cuadrados de los factores de escala 214 a través de un cuadro 202, 204; (2) normalizar la raíz cuadrada basándose en un número de canales presentes en el audio comprimido; y (3) comparar la raíz cuadrada normalizada con un umbral para determinar si el audio comprimido supera un nivel especificado. La raíz cuadrada normalizada de una suma de cuadrados de los factores de escala 214 proporciona una buena estimación del nivel de audio.

Tal función tiene utilidad, no como medio para medir con precisión el nivel de audio, sino como medio para determinar si todavía está presente audio. Aún cuando el nivel de audio medido es preciso hasta, quizás, solamente 5 dB, el presente invento puede determinar que está presente audio. Por tanto, si se determina que el nivel de audio para cierto número de paquetes secuenciales está sustancialmente por debajo de lo que sería de esperar normalmente (por ejemplo, más de 30 dB por debajo), entonces puede suponerse que algo ha fallado aguas arriba.

Para conseguir esta detección del nivel de audio, se realiza cierto número de adiciones a la lógica de la Fig. 3 anterior. Estas adiciones se describen en lo que sigue.

El bloque 320 utiliza una tabla para determinar un valor entero para cada factor de escala 214 correspondiente que represente el cuadrado del valor de pico de voltaje analógico derivado. El bloque 320 almacena la suma de estos cuadrados a través de un cuadro 202 o 204.

El bloque 322 obtiene la raíz cuadrada de la suma de los cuadrados almacenados en el bloque 320, en un punto en que el proceso 102 de alteración de ganancia ha completado su tratamiento de una trama 202 o 204. Se normaliza entonces la raíz cuadrada, dependiendo del número de canales presentes en el audio comprimido, que representa el cuadrado de la tensión de entrada estimada. La raíz cuadrada normalizada es comparada con un umbral para determinar si el audio comprimido supera un nivel especificado, por encima del cual un canal de audio puede ser declarado activo.

Además, la propia detección de nivel puede utilizarse para iniciar una alteración de los niveles de audio, formando por tanto un sencillo control de ganancia automático. Por ejemplo, si durante algún período de tiempo se ve que el nivel de audio es demasiado bajo o demasiado alto, entonces puede ajustarse el nivel de ganancia empleando la lógica de la Fig. 3, para llevar al nivel de audio a un valor predeterminado.

Esto lo ejecutarían los bloques 320 o 322 examinando el nivel de pico durante un cierto período de tiempo y, si se determina que el nivel es demasiado bajo o demasiado alto, alterando entonces la ganancia a un valor predeterminado mediante el empleo de la lógica de la Fig. 3. El examen del nivel de pico durante un largo período de tiempo mitiga los errores de medición y de control.

Ventajas

El presente invento incluye varias características y ventajas únicas:

- 1) La alteración del nivel de audio en una corriente de datos de audio MPEG debe realizarse sin un retardo apreciable. Generalmente, una descodificación y codificación de la corriente de datos de audio MPEG requiere, al menos, 48 ms de retardo. Sin embargo, para difusión, el audio está asociado con vídeo y, a no ser que se inyecte un retardo de vídeo adicional, a un espectador le parecerá que los labios se mueven mucho antes de que se oiga el diálogo, generándose un problema con la sincronización de los labios.
- 2) El presente invento demuestra una elevada eficacia desde el punto de vista del cálculo. Por ejemplo, un software de prueba ejecutado en un ordenador personal hizo variar el nivel de audio de una corriente de datos de audio MPEG a más de 20 veces el tiempo real, cuando una codificación y una descodificación MPEG solamente funcionaban en tiempo real.
- 3) Eliminación de los errores de cuantización del descodificador en el ínterin. En la técnica anterior si el descodificador solamente proporciona 16 bits de resolución, el propio descodificador podría inyectar errores de cuantización en la corriente de datos de audio MPEG. Esto es cierto si la corriente original de datos de audio MPEG estaba codificada con más de 16 bits de precisión (típicamente 20 o 24 bits). La mayoría de los descodificadores están contruidos para mantener, como máximo una precisión de 16 bits. Si se incrementa el nivel de audio tras una descodificación de 16 bits, el codificador siguiente ve una base de ruido elevada provocada por errores de truncamiento en el descodificador. Con este invento, si la codificación de audio MPEG original se realizó con más de 16 bits de precisión, la ganancia puede incrementarse mientras se mantiene la base de ruido en un descodificador de 16 bits a un nivel óptimo, incrementando realmente las relaciones entre señal y ruido.

Conclusión

5 La anterior descripción de la realización preferida del invento se ha presentado con fines de ilustración y descripción. No se pretende que sea exhaustiva ni que limite el invento a la forma precisa expuesta. A la luz de las enseñanzas anteriores son posibles muchas modificaciones y variaciones.

10 Por ejemplo, si bien la anterior exposición ofrece una realización del presente invento aplicada a un sistema de transmisión por satélite o a un reproductor MPEG personal, el presente invento puede utilizarse en cualquier aplicación que utilice audio MPEG. Además, aunque el presente invento se ha descrito en términos de audio MPEG, podría aplicarse, también, a otros esquemas de compresión, tales como Dolby[®] AC-3. Finalmente, aunque en esta memoria se ha descrito una lógica específica, los expertos en la técnica reconocerán que otra lógica puede ofrecer los mismos resultados, sin apartarse del alcance del presente invento.

15 Se pretende que el alcance del invento no esté limitado por esta descripción detallada, sino por las reivindicaciones adjuntas a ella. La anterior exposición, ejemplos y datos ofrecen una completa descripción de la fabricación y uso de la composición del invento. Dado que pueden ejecutarse muchas realizaciones del invento sin apartarse del alcance del mismo, el invento reside en las siguientes reivindicaciones adjuntas.

20

25

30

35

40

45

50

55

60

65

ES 2 315 992 T3

REIVINDICACIONES

5 1. Un método de detección del nivel de audio para audio comprimido (216) de una corriente (104) de datos, que comprende:

(a) extraer (320) factores de escala (214) para el audio comprimido (216) a partir de la corriente (104) de datos; y

10 (b) realizar (320, 322) una detección de nivel para el audio comprimido (216) utilizando los factores de escala (214) extraídos sin descomprimir el audio comprimido (216) a través de:

(1) determinar (320) el cuadrado de un valor de pico de voltaje analógico derivado para cada uno de los factores de escala (214);

15 (2) obtener (322) la raíz cuadrada de la suma de los cuadrados a través de un cuadro;

(3) normalizar (322) la raíz cuadrada en base a un número de canales presente en el audio comprimido (216); y

20 (4) comparar (322) la raíz cuadrada normalizada con un umbral para determinar si el audio comprimido (216) excede un nivel especificado.

25 2. Un aparato adaptado para llevar a cabo el método de la reivindicación 1.

30

35

40

45

50

55

60

65

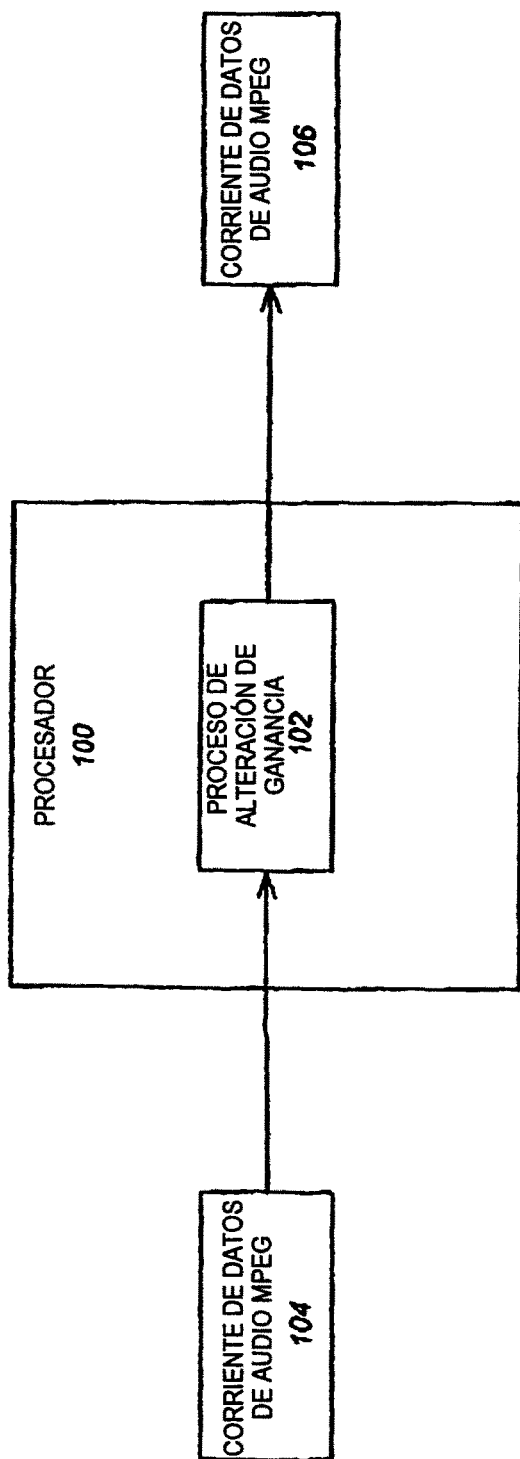


FIG. 1

CORRIENTE DE DATOS DE AUDIO MPEG

200



CABECERA 208	CRC 210	ASIGNACIÓN DE BIT 212	FACTORES DE ESCALA 214	MUESTRAS 216	DATOS AUXILIARES 218
-----------------	------------	-----------------------------	------------------------------	-----------------	----------------------------

CAPA I
202

CABECERA 208	CRC 210	ASIGNACIÓN DE BIT 212	SCFSI 220	FACTORES DE ESCALA 214	MUESTRAS 216	DATOS AUXILIARES 218
-----------------	------------	-----------------------------	--------------	------------------------------	-----------------	----------------------------

CAPA II
204

CABECERA 208	CRC 210	INFORMACIÓN LATERAL 222	DEPÓSITO DE BITS 224
-----------------	------------	-------------------------------	-------------------------

CAPA III
206

FIG. 2

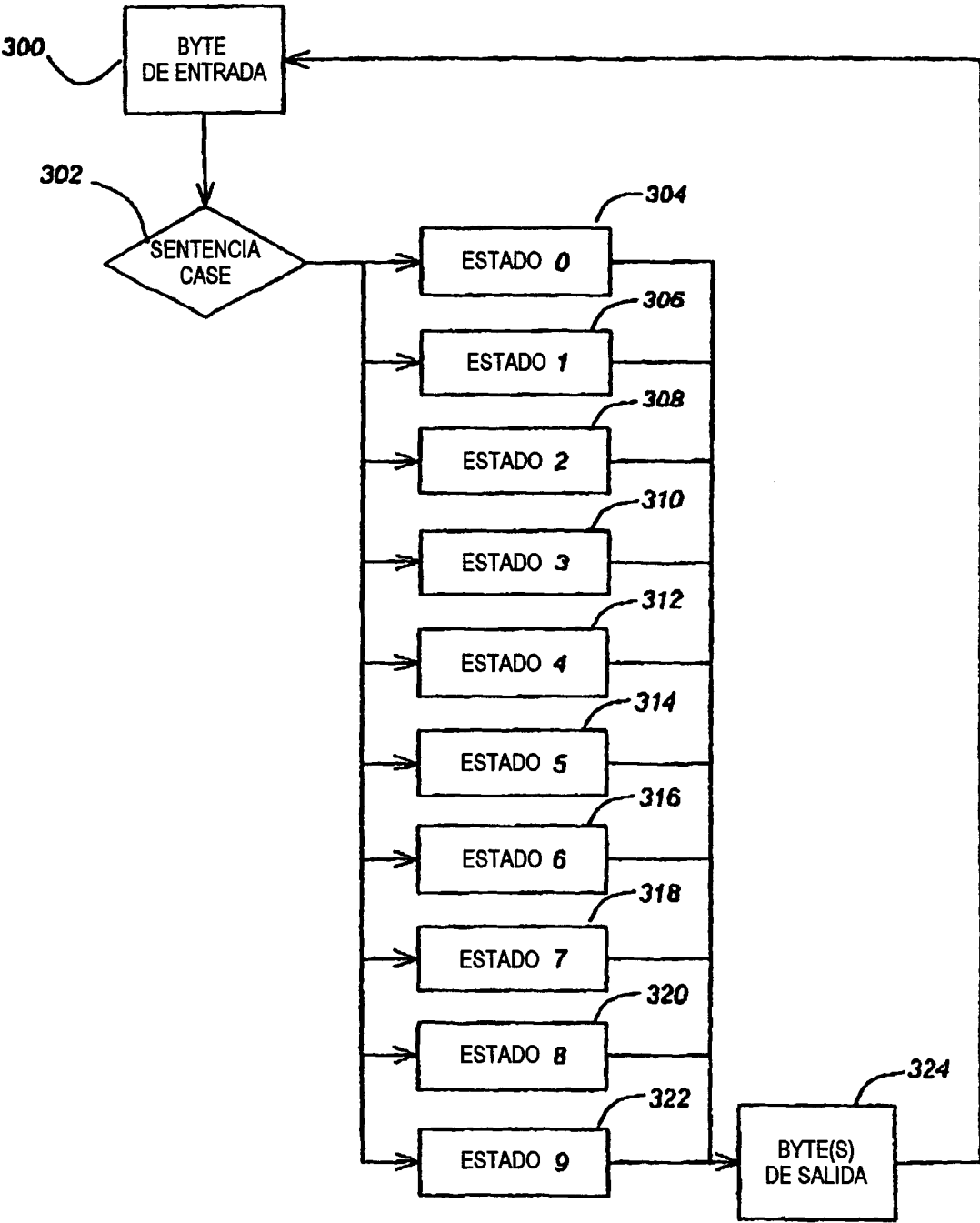


FIG. 3