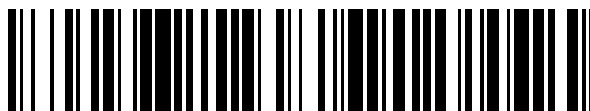


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 373 741**

51 Int. Cl.:
H04S 7/00 (2006.01)
G10L 11/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06739542 .6**
96 Fecha de presentación: **23.03.2006**
97 Número de publicación de la solicitud: **1878307**
97 Fecha de publicación de la solicitud: **16.01.2008**

54 Título: **MEDICIÓN ECONÓMICA DE LA INTENSIDAD DE UNA SEÑAL DE AUDIO CODIFICADA.**

30 Prioridad:
13.04.2005 US 671381 P

45 Fecha de publicación de la mención BOPI:
08.02.2012

45 Fecha de la publicación del folleto de la patente:
08.02.2012

73 Titular/es:
**DOLBY LABORATORIES LICENSING
CORPORATION
100 POTRERO AVENUE
SAN FRANCISCO, CALIFORNIA 94103-4813, US**

72 Inventor/es:
**CROCKETT, Brett, Graham;
SMITHERS, Michael, John y
SEEFELDT, Alan, Jeffrey**

74 Agente: **de Elzaburu Márquez, Alberto**

ES 2 373 741 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

Medición económica de la intensidad de una señal de audio codificada.

5 *Campo técnico*

La invención se refiere al procesamiento de señales de audio. Más particularmente, se refiere a un cálculo económico de una medición objetiva de la intensidad de una señal de audio codificada con baja tasa de bits, tales como una señal de audio codificada con Dolby Digital (AC-3), Dolby Digital Plus o Dolby E. "Dolby", "Dolby Digital", "Dolby Digital Plus" y "Dolby E" son marcas comerciales de Dolby Laboratories Licensing Corporation. Los aspectos de la invención también pueden ser usados en otros tipos de codificación de audio.

Antecedentes de la técnica

Los detalles de la codificación Dolby Digital se exponen en las referencias siguientes:

- 15 ATSC Standard A52/A: Digital Audio Compression Standard (AC-3), Revision A, Advanced Television Systems Committee, 20 Ag. 2001. El documento A/52A está disponible en la Web en la dirección <http://www.atsc.org/standards.html>.
Flexible Perceptual Coding for Audio Transmission and Storage," por Craig C. Todd, et al, 96ª Convención de la Audio Engineering Society, 26 Feb. 1994, Preprint 3796;
- 20 "Design and Implementation of AC-3 Coders," por Steve Vernon, IEEE Trans. Consumer Electronics, Vol. 41, No. 3, Ag. 1995.
"The AC-3 Multichannel Coder" by Mark Davis, Audio Engineering Society Preprint 3774, 95ª Convención AES, Octubre, 1993.
"High Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Applications," por Bosi et al, Audio Engineering Society Preprint 3365, 93ª Convención AES, Octubre, 1992.

Las patentes US Nos. 5.583.962, 5.632.005, 5.633.981, 5.727.119, 5.909.664 y 6.021.386.

- 30 Los detalles de la codificación Dolby Digital Plus se describen en "Introduction to Dolby Digital Plus, an Enhancement to the Dolby Digital Coding System," Artículo de la convención AES 6196, 117ª Convención AES, 28 Oct, 2004.

Los detalles de la codificación Dolby E se describen en "Efficient Bit Allocation, Quantization, and Coding in an Audio Distribution System", AES Preprint 5068, 107ª Conferencia AES, Agosto 1999 y "Professional Audio Coder Optimized for Use with Video", AES Preprint 5033, 107ª Conferencia AES Agosto 1999.

- 35 Una visión general de diversos codificadores perceptuales, incluyendo codificadores Dolby, codificadores MPEG, y otros, se describe en "Overview of MPEG Audio: Current and Future Standards for Low-Bit-Rate Audio Coding," por Karlheinz Brandenburg y Marina Bosi, J. Audio Eng. Soc., Vol. 45, No. 1/2, Enero/Febrero 1997.

- 40 Existen muchos procedimientos para medir objetivamente la intensidad percibida de las señales de audio. Los ejemplos de procedimientos incluyen mediciones de potencia ponderada (tales como LeqA, LeqB, LeqC), así como mediciones de intensidad basadas en psicoacústica, tales como "Acoustics – Method for Calculating Loudness Level", ISO532 (1975). Las mediciones ponderadas de intensidad de potencia procesan la señal de audio de entrada mediante la aplicación de un filtro predeterminado que enfatiza las frecuencias perceptiblemente más sensibles, mientras que resta importancia a las
- 45 frecuencias perceptiblemente menos sensibles y, a continuación, promedian la potencia de la señal filtrada durante un período de tiempo predeterminado. Los procedimientos psicoacústicos son, típicamente, más complejos y su objetivo es modelar mejor el funcionamiento del oído humano. Esto se consigue dividiendo la señal de audio en bandas de frecuencia que imitan la respuesta de frecuencia y la sensibilidad del oído y, a continuación, manipulan e integran estas bandas, teniendo en cuenta fenómenos psicoacústicos, tales como el enmascaramiento temporal y de frecuencia, así como la
- 50 percepción no lineal de la intensidad con intensidades diferentes de la señal. El objetivo de todos los procedimientos de medición objetiva de la intensidad es obtener una medición numérica de la intensidad que se ajuste estrechamente a la percepción subjetiva de la intensidad de una señal de audio.

- 55 La codificación perceptual o codificación de señal de audio a baja tasa de bits es usada, normalmente, para comprimir datos de señales de audio para un almacenamiento, una transmisión y un suministro eficiente en aplicaciones tales como emisión de televisión digital y la venta en línea de música por Internet. La codificación perceptual consigue su eficiencia transformando la señal de audio en un espacio de información, donde tanto las redundancias como los componentes de la señal que están enmascarados psicoacústicamente, pueden ser descartados fácilmente. La información restante es empaquetada en una secuencia o en un archivo de información digital. Típicamente, la medición de la intensidad de la
- 60 señal de audio, representada por una señal de audio codificada a baja tasa de bits, requiere decodificar de nuevo la señal de audio al dominio temporal (por ejemplo, PCM), lo que puede ser computacionalmente intensivo. Sin embargo, algunas

señales codificadas perceptualmente, a baja tasa de bits, contienen información que puede ser útil para un procedimiento de medición de intensidad, ahorrando, de esta manera, el costo computacional de decodificar totalmente la señal de audio. Dolby Digital (AC-3), Dolby Digital Plus y Dolby E se encuentran entre estos sistemas de codificación de audio.

5 Los codificadores de audio perceptuales, de baja tasa de bits, Dolby Digital, Dolby Digital Plus y Dolby E dividen las señales de audio en segmentos de ventanas de tiempo, superpuestas (o bloques de codificación de audio), que son transformados a una representación en el dominio de la frecuencia. La representación en el dominio de la frecuencia de los coeficientes espectrales es expresada mediante una notación exponencial que comprende conjuntos de un exponente y la mantisa asociada. Los exponentes, que funcionan como factores de escala, son empaquetados en la secuencia de audio codificada. Las mantisas representan los coeficientes espectrales después de haber sido normalizadas por los exponentes. A continuación, los exponentes son pasados a través de un modelo perceptual del oído y usados para cuantificar y empaquetar las mantisas en la secuencia de audio codificada. Tras la decodificación, los exponentes son desempaquetados desde la secuencia de audio codificada y, a continuación, son pasados a través del mismo modelo perceptual para determinar cómo desempaquetar las mantisas. A continuación, las mantisas son desempaquetadas, son combinadas con los exponentes para crear una representación en el dominio de la frecuencia de la señal de audio que, a continuación, es decodificada y convertida de nuevo a una representación en el dominio temporal.

Debido a que muchas mediciones de intensidad incluyen cálculos de potencia y del espectro de potencia, el ahorro computacional puede conseguirse decodificando solo parcialmente la señal de audio codificada a baja tasa de bits y pasando la información decodificada parcialmente (tal como el espectro de potencia) a la medición de intensidad. La invención es útil siempre que haya una necesidad de medir la intensidad, pero no de decodificar la señal de audio. Toma ventaja del hecho de que una medición de intensidad puede hacer uso de una versión aproximada de la señal de audio, no siendo dicha aproximación, normalmente, adecuada para ser escuchada. Un aspecto de la presente invención es el reconocimiento de que una representación tosca de la señal de audio, que está disponible, sin descodificar completamente una secuencia de bits, en muchos sistemas de codificación de audio, puede proporcionar una aproximación del espectro de audio que puede ser utilizada en la medición de la intensidad de la señal de audio. En la codificación de audio Dolby Digital, Dolby Digital Plus y Dolby E, los exponentes proporcionan una aproximación del espectro de potencia de la señal de audio. De manera similar, en ciertos sistemas de codificación diferentes, los factores de escala, envolventes espectrales y los coeficientes predictivos lineales pueden proporcionar una aproximación del espectro de potencia de la señal de audio. Estos y otros aspectos y ventajas de la invención se comprenderán mejor con la lectura y la comprensión del resumen y la descripción siguientes de la invención.

El documento US 2001/0027393 A1 divulga un sistema de audio conferencia compuesto por N terminales conectados respectivamente a una unidad de control multipunto. Cada terminal comprende un codificador cuya entrada recibe datos de audio a transmitir a los otros terminales y cuya salida está conectada a una entrada de la unidad de control multipunto. Cada terminal tiene también un decodificador cuya entrada está conectada a una salida de la unidad de control multipunto y cuya salida suministra datos que son transmitidos al terminal considerado por los otros terminales. La unidad de control multipunto está compuesta, esencialmente, de un combinador que combina las señales presentes en sus entradas y suministra a la entrada del decodificador de un terminal una señal que representa la suma de las señales suministradas respectivamente por todos los codificadores de los N terminales, excepto para la señal de ese terminal. La unidad de control multipunto tiene también N decodificadores parciales destinados a recibir, respectivamente, las tramas de audio producidas por los N terminales para decodificarlas y suministrar las mismas, de esta manera, a las entradas del combinador. La unidad de control multipunto tiene N decodificadores parciales que tienen salidas conectadas, respectivamente, a las entradas de los decodificadores de los terminales y que tienen entradas conectadas a las salidas del combinador. El documento describe el cálculo de la energía total de todos los terminales, menos uno, en cada banda de frecuencia.

Un objeto de la invención es proporcionar una medición computacionalmente económica de la intensidad percibida de la señal de audio codificada en una baja tasa de bits.

Este objeto se consigue mediante el procedimiento según la reivindicación 1, el aparato según la reivindicación 12 y un programa de ordenador almacenado en un medio que puede ser leído por ordenador según la reivindicación 24, respectivamente. Las realizaciones preferentes de la invención se definen en las reivindicaciones dependientes.

55 La invención consigue el objeto según se reivindica en las reivindicaciones independientes, decodificando solo parcialmente el material de audio y pasando la información parcialmente decodificada a una medición de intensidad. El procedimiento aprovecha las propiedades específicas de la información de audio parcialmente decodificada, tales como los exponentes en la codificación de audio Dolby Digital, Dolby Digital Plus y Dolby E.

60 Un primer aspecto de la invención mide la intensidad de una señal de audio codificada en una secuencia de bits que incluye datos a partir de los cuales puede derivarse una aproximación del espectro de potencia de la señal de audio, sin

decodificar completamente la señal de audio, derivando la aproximación del espectro de potencia de la señal de audio a partir de la secuencia de bits sin decodificar completamente la señal de audio, y determinando una intensidad aproximada de la señal de audio en respuesta a la aproximación del espectro de potencia de la señal de audio.

5 En este primer aspecto de la invención, los datos incluyen representaciones toscas de la señal de audio y representaciones más finas asociadas de la señal de audio, y la aproximación del espectro de potencia de la señal de audio es derivada a partir de las representaciones toscas de la señal de audio.

10 En un aspecto adicional de la invención, la señal de audio codificada en una secuencia de bits puede ser una señal de audio codificada en sub-bandas, que tiene una pluralidad de sub-bandas de frecuencia, teniendo cada sub-banda un factor de escala y datos de muestra asociados con el mismo, y en el que las representaciones toscas de la señal de audio comprenden factores de escala y las representaciones más finas asociadas de la señal de audio comprenden datos de muestra asociados con cada factor de escala.

15 En todavía un aspecto adicional de la invención, el factor de escala y los datos de muestra de cada sub-banda pueden representar coeficientes espectrales en la sub-banda, por medio de una notación exponencial en la que el factor de escala comprende un exponente y los datos de muestra asociados comprenden mantisas.

20 En todavía un aspecto adicional de la invención, la señal de audio codificada en una secuencia de bits puede ser una señal de audio con una codificación predictiva lineal, en la que las representaciones toscas de la señal de audio comprenden los coeficientes predictivos lineales y las representaciones más finas de la señal de audio comprenden información de excitación asociada con los coeficientes predictivos lineales.

25 En todavía un aspecto adicional de la invención, las representaciones toscas de la señal de audio pueden comprender al menos una envolvente espectral y las representaciones más finas de la señal de audio pueden comprender componentes espectrales asociados con al menos una envolvente espectral.

30 En todavía otro aspecto adicional de la invención, la determinación de una intensidad aproximada de la señal de audio, en respuesta a la aproximación del espectro de potencia de la señal de audio, puede incluir la aplicación de una medición ponderada de la intensidad de potencia. La medición ponderada de la intensidad de potencia puede emplear un filtro que resta importancia a las frecuencias menos audibles y calcula un promedio en el tiempo de la potencia de la señal de audio filtrada.

35 En todavía otro aspecto de la invención, la determinación de una intensidad aproximada de la señal de audio, en respuesta a la aproximación del espectro de potencia de la señal de audio, puede incluir la aplicación de una medición de intensidad psicoacústica. La medición de intensidad psicoacústica puede emplear un modelo del oído humano para determinar la intensidad específica en cada una de entre una pluralidad de bandas de frecuencia, similares a las bandas críticas del oído humano. En un entorno de codificador de subbandas, las sub-bandas pueden ser similares a las bandas críticas del oído humano y la medición de intensidad psicoacústica puede emplear un modelo del oído humano para determinar una intensidad específica en cada una de las sub-bandas.

40 Los aspectos de la invención incluyen procedimientos que ponen en práctica las funciones anteriores, medios que ponen en práctica las funciones, aparatos que ponen en práctica los procedimientos y un programa de ordenador, almacenado en un medio que puede ser leído por ordenador, para hacer que un ordenador realice los procedimientos que ponen en práctica las funciones anteriores.

Descripción de los dibujos

La Fig. 1 muestra un diagrama de bloques funcional, esquemático, de una disposición general para la medición de la intensidad de una señal de audio codificada a baja tasa de bits.

50 La Fig. 2 muestra un diagrama de bloques funcional, esquemático, generalizado, de un decodificador Dolby Digital, Dolby Digital Plus y Dolby E.

Las Figs. 3a y 3b muestran diagramas de bloques funcionales, esquemáticos, de dos disposiciones generales para el cálculo de una medición de intensidad objetiva usando la potencia ponderada y mediciones basadas en psicoacústica, respectivamente.

55 La Fig. 4 muestra las ponderaciones de frecuencias comunes usadas en la medición de intensidad según la disposición del ejemplo de la Fig. 3a.

La Fig. 5 es un diagrama de bloques, esquemático, funcional, que muestra una disposición general más económica para la medición de la intensidad de la señal de audio codificada según los aspectos de la invención.

60 Las Figs. 6a y 6b son diagramas de bloques, esquemáticos, funcionales, de la disposición más económica para la medición de la intensidad, que incorporan las disposiciones de intensidad mostradas en los ejemplos de las Figs. 3a y 3b, según los aspectos de la invención.

Mejor modo de llevar a cabo la invención

Uno de los beneficios de los aspectos de la presente invención es la medición de la intensidad de audio codificada a baja tasa de bits, sin necesidad de decodificar completamente la señal de audio a PCM, cuya decodificación incluye etapas de procesamiento de decodificación caras, tales como la asignación de bits, de-cuantificación, una transformación inversa, etc. Los aspectos de la invención reducen de manera importante los requisitos de procesamiento (sobrecarga computacional). Este enfoque es beneficioso cuando se desea una medición de intensidad, pero no se necesita la señal de audio decodificada.

Los aspectos de la presente invención pueden ser usados, por ejemplo, en entornos tales como los divulgados en (1) la solicitud de patente US, no provisional, en trámite, 11/373.577 y la publicación No. 200600002572, presentada el 1 de Julio de 2004 y publicada el 5 de Enero de 2006, titulada "Method for Correcting Metadata Affecting the Playback Loudness and Dynamic Range of Audio Information", por Smithers et al., y (2) en la realización de una medición de intensidad y corrección en un almacenamiento compartido o en una cadena de transmisión en la que el acceso al audio decodificado no se necesita y no es deseable.

El ahorro de procesamiento proporcionado por los aspectos de la invención contribuyen también a hacer posible la realización de una medición de intensidad y una corrección de metadatos (por ejemplo, cambiar un parámetro DIALNORM al valor correcto) en tiempo real en un gran número de datos de señales de audio con una compresión de datos a baja tasa de bits. Frecuentemente, muchas señales codificadas a baja tasa de bits son multiplexadas y transportadas en flujos de transporte MPEG. La medición de intensidad según los aspectos de la presente invención realiza mediciones de intensidad en tiempo real en un gran número de señales de audio comprimidas, de manera mucho más fiable, en comparación con los requisitos de decodificar completamente las señales de audio comprimidas a PCM para realizar la medición de intensidad.

La Fig. 1 muestra una disposición 100 de la técnica anterior para medir la intensidad de una señal de audio codificada. La información o los datos 101 de audio digital codificado, tales como una señal de audio que ha sido codificada a baja tasa de bits, es decodificada por un decodificador o una función de decodificación ("Decodifica") 102, por ejemplo, en una señal de audio PCM 103. A continuación, esta señal es aplicada a un medidor de intensidad o a un procedimiento o un algoritmo de medición ("Mide Intensidad") 104, que genera un valor 105 de intensidad medida.

La Fig. 2 muestra un diagrama de bloques 200, estructural o funcional, de la técnica anterior, de una etapa Decodifica 102. La estructura o las funciones que muestra son representativas de los decodificadores Dolby Digital, Dolby Digital Plus y Dolby E. Las tramas de datos de audio codificado 101 son aplicadas a una función desempaquetadora o a un desempaquetador de datos ("Sinc. Tramas, detección de errores y Eliminación de Formato de Tramas") 202, que desempaqueta los datos aplicados en los datos de exponentes 203, datos de mantisas 204 y otra información variada de asignación de bits 207. Los datos de exponentes 203 son convertidos a un espectro de potencia logarítmico 206 por un dispositivo o una función ("Espectro Potencia Log") 205, y este espectro de potencia logarítmico es usado por un asignador de bits o una función de asignación de bits ("Asignación de Bits") 208, para calcular la señal 209, que es la longitud, en bits, de cada mantisa cuantificada. A continuación, las mantisas son de-cuantificadas y son combinadas con los exponentes por un dispositivo o una función ("De-cuantifica Mantisas") 210, para proporcionar una salida 211 y son convertidas de nuevo al dominio temporal mediante una función o un dispositivo de banco de filtros inverso ("Banco de Filtros inverso") 212. El banco de filtros inverso 212 también superpone y suma una porción del resultado del banco de filtros inverso actual con el resultado del banco de filtros inverso anterior (en el tiempo) para crear la señal de audio decodificada 103. En las implementaciones prácticas del decodificador, los dispositivos o funciones Asignación de Bits, De-Cuantifica Mantisas y Banco de Filtros Inverso requieren considerables recursos de computación. Pueden encontrarse más detalles sobre el proceso de decodificación en algunas de las referencias indicadas anteriormente.

Las Figs. 3a y 3b muestran disposiciones de la técnica para medir objetivamente la intensidad de una señal de audio. Estas representan variaciones de la etapa Mide Intensidad 104 (Fig 1). Aunque las Figs. 3a y 3b muestran, respectivamente, ejemplos de dos categorías generales de técnicas de medición objetiva de intensidad, la elección de una técnica objetiva de medición particular no es crítica para la invención y pueden emplearse otras técnicas de medición objetiva de intensidad.

La Fig. 3a muestra un ejemplo de la medición ponderada de potencia 300 usada normalmente en la medición de la intensidad. Una señal de audio 103 es pasada a través de un filtro de ponderación o una función de filtrado ("Filtro de Ponderación") 302, que está diseñado para enfatizar las frecuencias perceptiblemente más sensibles, mientras que resta importancia a las frecuencias perceptiblemente menos sensibles. La potencia 305 de la señal filtrada 303 es calculada mediante un dispositivo o una función ("Potencia") 304 y es promediada durante un período de tiempo definido por un dispositivo o una función ("Promedio") 306, para crear un valor de intensidad 105. Existen una serie de diferentes características de un filtro de ponderación estándar y algunos ejemplos normales se muestran en la Fig. 4. En la práctica,

se usan, frecuentemente, versiones modificadas de la disposición de la Fig. 3a, en las que las modificaciones, por ejemplo, previenen que períodos de tiempo de silencio sean incluidos en el promedio.

5 Frecuentemente, se usan también técnicas basadas en psicoacústica para medir la intensidad. La Fig. 3b muestra una
 10 disposición típica 310 de la técnica anterior de dicha una disposición basada en psicoacústica. Una señal de audio 103 es
 filtrada por un filtro de transmisión o una función de filtrado ("Filtro de Transmisión") 312, que representa la respuesta en
 magnitud a frecuencia variable del oído externo y medio. A continuación, la señal filtrada 313 es separada por un banco
 15 de filtros auditivo o una función banco de filtros ("Banco de filtros auditivo") 314, en bandas de frecuencia 315 que son
 equivalentes a, o más estrechas que, las bandas auditivas críticas. Esto puede conseguirse realizando una transformada
 rápida de Fourier (FFT) (tal como se implementa, por ejemplo, por medio de una transformada de frecuencia discreta
 (DFT)) y, a continuación, agrupando las bandas linealmente espaciadas en bandas que se aproximan a las bandas
 20 críticas del oído (tal como en una escala Barck o ERB). Como alternativa, esto puede conseguirse por medio de un único
 filtro pasa banda para cada banda Barck o ERB. A continuación, cada banda es convertida por un dispositivo o una
 función ("Excitación") 316 en una señal de excitación 317, que representa la cantidad de estímulos o de excitación
 experimentada por el oído humano dentro de la banda. La intensidad percibida o intensidad específica para cada banda
 319 es calculada, a continuación, a partir de la excitación por un dispositivo o una función ("Intensidad Específica") 318, y
 la intensidad específica a lo largo de todas las bandas es sumada por un sumador o una función sumadora ("Suma") 320,
 para crear una única medición de intensidad 105. El procedimiento de suma puede tener en consideración varios efectos
 perceptuales, por ejemplo, el enmascaramiento de frecuencia. En implementaciones prácticas de estos procedimientos
 25 perceptuales, se necesitan considerables recursos computacionales para el filtro de transmisión y el banco de filtros
 auditivo.

La Fig. 5 muestra un diagrama de bloques 500 de un aspecto de la presente invención. Una señal de audio digital
 25 codificada 101 es parcialmente decodificada por un dispositivo o una función ("Decodificación Parcial") 502 y la intensidad
 es medida a partir de la información parcialmente decodificada 503 por un dispositivo o una función ("Mide Intensidad")
 504. Dependiendo de cómo se realice la decodificación parcial, la medición de la intensidad resultante 505 puede ser muy
 similar a, pero no exactamente igual que, la medición de intensidad 105 calculada a partir de la señal de audio
 completamente decodificada 103 (Fig. 1). En el contexto de las implementaciones Dolby Digital, Dolby Digital Plus y Dolby
 E de los aspectos de la invención, una decodificación parcial puede incluir la omisión de los dispositivos o las funciones
 30 Asignación de Bits, De-cuantificación de Mantisas y Banco de Filtro Inverso de un decodificador, tal como el ejemplo de la
 Fig. 2.

Las Figs. 6a y 6b muestran dos ejemplos de implementaciones de la disposición general de la Fig. 5. Aunque ambos
 35 pueden emplear el mismo dispositivo o función Decodifica Parcial 502, cada uno puede tener un dispositivo o una función
 Mide Intensidad 504 diferente, siendo el del ejemplo de la Fig 6a 600 similar al del ejemplo de la Fig. 3A, y siendo el del
 ejemplo de la Fig. 6B similar al del ejemplo de la Fig. 3b. En ambos ejemplos, Decodifica Parcial 502 extrae sólo los
 exponentes 203 de la secuencia de audio codificada y convierte los exponentes a un espectro de potencia 206. Dicha
 extracción, puede ser realizada por un dispositivo o una función ("Sin. Trama, Detección de Errores y Eliminación de
 Formato de Trama") 202, tal como en el ejemplo de la Fig. 2, y dicha conversión puede ser realizada por un dispositivo o
 40 una función ("Espectro de Potencia Log") 205, tal como en el ejemplo de la Fig. 2. No hay ningún requisito para de-
 cuantificar las mantisas, realizar una asignación de bits, y realizar un banco de filtros inverso, tal como se requeriría para
 una decodificación completa, tal como se muestra en el ejemplo de decodificación de la Fig. 2.

El ejemplo de la Fig. 6a incluye una función Mide Intensidad 504, que puede ser una versión modificada del medidor de
 45 intensidad o de la función de medición de intensidad de la Fig. 3a. En este ejemplo, se aplica un filtrado de ponderación
 modificado en el dominio de la frecuencia, aumentando o disminuyendo los valores de potencia en cada banda por un
 filtro de ponderación o una función de filtrado ponderado ("Filtro de ponderación Modificado") 601. Por el contrario, el
 ejemplo de la Fig. 3a aplica un filtrado de ponderación en el dominio temporal. Aunque opera en el dominio de la
 frecuencia, el Filtro de Ponderación Modificado afecta al audio de la misma manera que el Filtro de Ponderación en el
 50 dominio temporal de la Fig. 3a. El filtro 601 está "modificado" con respecto al filtro 302 de la Fig. 3a en el sentido de que
 opera en valores de amplitud logarítmica en lugar de valores lineales, y que opera sobre una escala de frecuencias no
 lineal, en lugar de en una escala de frecuencias lineal. El espectro de frecuencia de potencia ponderada 602 es convertido,
 a continuación, a potencia lineal y se suma en frecuencia y se promedia en el tiempo por un dispositivo o una función
 ("Convierte, Suma y Promedia") 603, aplicando, por ejemplo, la ecuación 5, a continuación. La salida es un valor de
 55 intensidad objetivo 505.

El ejemplo de la Fig. 6b incluye una función Mide Intensidad 504, que puede ser una versión modificada del medidor de
 intensidad o la función de medición de intensidad de la Fig. 3b. En este ejemplo, un filtro de transmisión modificado o una
 60 función de filtrado ("Filtro de Transmisión Modificado") 611 es aplicado directamente en el dominio de la frecuencia,
 aumentando o disminuyendo los valores de potencia logarítmica en cada banda. Por el contrario, el ejemplo de la Fig. 3b
 aplica un filtrado de ponderación en el dominio temporal. Aunque opera en el dominio de la frecuencia, el Filtro de

Transmisión Modificado afecta al audio en la misma manera que el Filtro de Transmisión en el dominio temporal de la Fig. 3b. Un banco de filtros auditivos modificado o una función Banco de Filtros ("Banco de Filtros Auditivo Modificado") 613 acepta como entrada el espectro de potencia logarítmica, bandas de frecuencia espaciadas linealmente y divide o combina estas bandas linealmente espaciadas en una salida 315 del banco de filtros de bandas críticas espaciadas (por ejemplo, bandas Bark o ERB). El banco de Filtros Auditivo Modificado 613 convierte también la señal de potencia en el dominio logarítmico en una señal lineal para el dispositivo o la función de excitación siguiente ("Excitación") 316. El Banco de Filtros Auditivo Modificado 613 está "modificado" con respecto al Banco de Filtros Auditivo 314 de la Fig. 3b en el sentido de que opera sobre valores de amplitud logarítmica en lugar de sobre valores lineales y convierte dichos valores de amplitud logarítmica en valores lineales. Como alternativa, la agrupación de bandas en las bandas Bark o ERB puede realizarse en el Banco de Filtros Modificado 613, en lugar de en el Filtro de Transmisión Modificado 611. El ejemplo de la Fig. 6b incluye también una Intensidad Específica 318 para cada banda y una Suma 320, tal como en el ejemplo de la Fig. 3b.

Para las disposiciones mostradas en las Figs. 6a y 6b, se consiguen ahorros computacionales considerables ya que la decodificación no requiere una asignación de bits, una de-cuantificación de mantisas y un Banco de Filtros inverso. Sin embargo, para ambas disposiciones de la Fig. 6a y la Fig. 6b, la medición objetiva de intensidad resultante puede que no sea exactamente la misma que la medición calculada a partir de una señal de audio decodificada totalmente. Esto se debe a que parte de la información es descartada y, de esta manera, la información de audio usada para la medición está incompleta. Cuando los aspectos de la presente invención son aplicados a Dolby Digital, Dolby Digital Plus o Dolby E, la información de la mantisa es descartada y sólo se retienen los valores de exponentes cuantificados toscamente. Para Dolby Digital y Dolby Digital Plus, los valores están cuantificados en incrementos de 6 dB y para Dolby E están cuantificados en incrementos de 3 dB. Las menores etapas de cuantificación en Dolby E resultan en valores de exponentes cuantificados más finamente y, consiguientemente, una estimación más precisa del espectro de potencia.

Frecuentemente, los codificadores perceptuales se diseñan para alterar la longitud de los segmentos de tiempo superpuestos, denominada también tamaño de bloque, en conjunción con ciertas características de la señal de audio. Por ejemplo, Dolby Digital usa dos tamaños de bloque, un bloque más grande de 512 muestras, principalmente para señales de audio estacionarias, y un bloque más corto de 256 muestras, para señales de audio más transitorias. El resultado es que el número correspondiente de valores de espectro de potencia logarítmica 206 varía bloque a bloque. Cuando el tamaño del bloque es de 512 muestras, hay 256 bandas, y cuando el tamaño del bloque es de 256 muestras, hay 128 bandas.

Hay muchas maneras en las que los procedimientos propuestos en las Figs. 6a y 6b pueden manipular tamaños de bloque variables y cada manera conduce a una medición de intensidad resultante similar. Por ejemplo, el Espectro de Potencia Logarítmica 205 puede ser modificado para sacar siempre un número de bandas constante a una tasa de bloques constante, combinando o promediando múltiples bloques más pequeños en bloques más grandes y esparciendo la potencia desde el menor número de bandas al mayor número de bandas. Como alternativa, Mide Intensidad puede aceptar tamaños de bloque variables y ajustar, consiguientemente, sus procedimientos de filtrado, excitación, intensidad específica, promedio y suma, por ejemplo, ajustando las constantes de tiempo.

Ejemplo de medición de potencia ponderada

Como ejemplo de los aspectos de la presente invención, una versión altamente económica de un procedimiento de medición de intensidad de potencia ponderada puede usar secuencias de bits de Dolby Digital y la medición de intensidad de potencia ponderada LeqA. En este ejemplo altamente económico, sólo los exponentes cuantificados contenidos en una secuencia de bits Dolby Digital son usados como una estimación del espectro de la señal de audio, para realizar la medición de intensidad. Esto evita los requisitos computacionales adicionales de realizar una asignación de bits para recrear la información de la mantisa, que, por otro lado, solo proporcionaría una estimación ligeramente más precisa del espectro de la señal.

Tal como se muestra en los ejemplos de las Figs. 5 y 6a, la secuencia de bits Dolby Digital es decodificada parcialmente para recrear y extraer el espectro de potencia logarítmico, calculado a partir de los datos de exponentes cuantificados contenidos en la secuencia de bits. Dolby Digital realiza una codificación de audio a baja tasa de bits, enmarcando en una ventana 512 muestras de audio PCM consecutivas, superpuestas al 50%, y realizando una transformada TCMD, resultando en 256 coeficientes TCMD que son usados para crear la secuencia de audio codificada a baja tasa de bits. La decodificación parcial realizada en las Figs. 5 y 6a, desempaqueta los datos de exponentes $E(k)$ y convierte los datos desempaquetados a 256 valores de espectro de potencia logarítmica cuantificados, $P(k)$, que forman una representación espectral tosca de la señal de audio. Los valores espectrales de la potencia logarítmica, $P(k)$, están en unidades de dB. La conversión es tal como se indica a continuación

$$P(k) = -E(k) \cdot 20 \cdot \log_{10}(2) \quad 0 \leq k < N \quad (1)$$

donde $N = 256$, el número de coeficientes de transformación para cada bloque en una secuencia de bits Dolby Digital. Para usar el espectro de potencia logarítmico en el cálculo de la medición de potencia ponderada de intensidad, el espectro de potencia logarítmico es ponderado usando una curva de intensidad adecuada, tal como una de las curvas de ponderación A, B o C que se muestran en la Fig. 4. En este caso, se está calculando la medición de potencia LeqA y, por lo tanto, la curva de ponderación A es apropiada. Los valores del espectro de potencia logarítmico $P(k)$ son ponderados sumándolos a valores discretos de frecuencia de ponderación A, $A_W(k)$, también en unidades de dB

$$P_W(k) = P(k) + A_W(k) \quad 0 \leq k < N \quad (2)$$

Los valores discretos de frecuencia de ponderación A, $A_W(k)$, son creados calculando de los valores de ganancia de ponderación A para las frecuencias discretas, $f_{discreta}$, donde

$$f_{discreta} = \frac{F}{2} + F \cdot k \quad 0 \leq k < N \quad (3)$$

donde

$$F = \frac{F_s}{2 \cdot N} \quad 0 \leq k < N \quad (4)$$

y donde la frecuencia de muestreo F_s es, típicamente, igual a 48 kHz para Dolby Digital. Cada conjunto de valores del espectro de potencia logarítmica ponderada, $P_W(k)$, es convertido, a continuación, de dB a potencia lineal y se suma para crear la estimación de potencia ponderada A, P_{POW} , de las 512 muestras de audio PCM como

$$P_{POW} = \sum_{k=0}^{N-1} 10^{(P_W(k)/10)} \quad (5)$$

Tal como se ha indicado anteriormente, cada secuencia de bits Dolby Digital contiene transformadas consecutivas creadas creando ventanas de 512 muestras PCM con un solapamiento del 50% y realizando la transformada MDCT. Por lo tanto, una aproximación de la potencia ponderada A total, P_{TOT} , de la señal de audio codificada a baja tasa de bits en una secuencia de bits Dolby Digital puede ser calculada promediando los valores de potencia en todas las transformaciones en la secuencia de bits Dolby Digital, tal como se indica a continuación

$$P_{TOT} = \frac{1}{M} \sum_{m=0}^{M-1} P_{POW}(m) \quad (6)$$

donde M es igual al número total de transformadas contenidas en la secuencia de bits Dolby Digital. A continuación, la potencia media es convertida a unidades de dB, tal como se indica a continuación

$$L_A = 10 \cdot \log_{10}(P_{TOT}) - C \quad (7)$$

donde C es un desplazamiento constante debido a los cambios de nivel realizados en el proceso de transformación durante la codificación de la secuencia de bits Dolby Digital.

Ejemplo de medición psicoacústica

Como otro ejemplo de los aspectos de la presente invención, una versión altamente económica de un procedimiento de medición de intensidad de potencia ponderada puede usar secuencias de bits Dolby Digital y una medición de intensidad psicoacústica. En este ejemplo altamente económico, al igual que en el anterior, sólo los exponentes cuantificados contenidos en una secuencia de bits Dolby Digital son usados como una estimación del espectro de la señal de audio

para realizar la medición de intensidad. Al igual que en el otro ejemplo, esto evita los requisitos adicionales de realizar una asignación de bits para recrear la información de la mantisa, que, por otro lado, sólo proporcionaría una estimación ligeramente más precisa del espectro de la señal.

5 La solicitud de patente internacional No. PCT/US2004/016964, presentada el 27 de mayo 2004, Seefeldt et al, publicada como WO 2004/111994 A2 el 23 de diciembre de 2004, cuya solicitud designa los Estados Unidos, divulga, entre otras cosas, una medición objetiva de la intensidad percibida en base a un modelo psicoacústico. Los valores del espectro de potencia logarítmica, $P(k)$, derivados a partir de la decodificación parcial de una secuencia de bits Dolby Digital, pueden servir como entradas para una técnica, tal como en dicha solicitud internacional, así como otras mediciones psicoacústicas similares, en lugar de la señal de audio PCM original. Dicha disposición se muestra en el ejemplo de la Fig. 6b. Tomando prestadas la terminología y la notación de dicha solicitud PCT, una señal de excitación $E(b)$, que aproxima la distribución de energía a lo largo de la membrana basilar del oído interno, en la banda crítica B, puede ser aproximada a partir de los valores del espectro de potencia logarítmica, tal como se indica a continuación:

$$15 \quad E(b) = \sum_k |T(k)|^2 |H_b(k)|^2 10^{P(k)/10} \quad (8)$$

20 donde $T(k)$ representa la respuesta de frecuencia del filtro de la transmisión y $H_b(k)$ representa la respuesta de frecuencia de la membrana basilar en una posición que corresponde a la banda crítica B, siendo muestreadas ambas respuestas a la frecuencia correspondiente para transformar bin k. A continuación, las excitaciones correspondientes a todas las transformadas en la secuencia de bits Dolby Digital son promediadas, para producir una excitación total:

$$25 \quad \bar{E}(b) = \frac{1}{M} \sum_m E(b, m) \quad (9)$$

30 Usando las curvas de igual nivel intensidad, la excitación total en cada banda es transformada en un nivel de excitación que genera la misma intensidad a 1 kHz. La intensidad específica, una medición de la intensidad distribuida a través de la frecuencia, es calculada, a continuación, a partir de la excitación transformada, $E_{1kHz}(b)$, por medio de una no-linealidad compresiva:

$$35 \quad N(b) = G \left(\left(\frac{\bar{E}_{1kHz}(b)}{TQ_{1kHz}} \right)^\alpha - 1 \right) \quad (10)$$

40 donde TQ_{1kHz} es el umbral en silencio a 1 kHz y las constantes G y α son elegidas para ajustar los datos generados a partir de experimentos psicoacústicos que describen el crecimiento de la intensidad. Por último, la intensidad total, L, representada en unidades sone, es calculada sumando la intensidad específica en las bandas:

$$45 \quad L = \sum_b N(b) \quad (11)$$

50 Para los propósitos de ajustar la señal de audio, puede ser deseable calcular una ganancia de ajuste, G_{ajuste} , que cuando es multiplicada por la señal de audio hace que la intensidad de la señal de audio ajustada sea igual a cierta intensidad de referencia, I_{REF} , medida por medio de la técnica psicoacústica descrita. Debido a que la medición psicoacústica implica una no-linealidad en el cálculo de la intensidad específica, no existe una solución en forma cerrada para G_{ajuste} . Por el contrario, puede emplearse una técnica interactiva descrita en dicha solicitud PCT, en la que el cuadrado de la ganancia de ajuste es ajustada y multiplicada por la excitación total, $E(b)$, hasta que la intensidad total correspondiente, L, se encuentre dentro de una diferencia de umbral con respecto a la intensidad de referencia, I_{REF} . Entonces, la intensidad de la señal de audio se puede expresar en dB con respecto a la referencia, como:

$$60 \quad L_{dB} = 20 \log_{10} \left(\frac{1}{G_{ajuste}} \right) \quad (12)$$

Otros codecs de audio perceptuales

5 Los aspectos de la presente invención no se limitan a los sistemas de codificación Dolby Digital, Dolby Digital Plus y Dolby E. Las señales de audio codificadas usando ciertos otros sistemas de codificación diferentes, en los que una aproximación del espectro de potencia de audio es proporcionada, por ejemplo, por factores de escala, envolventes espectrales y coeficientes predictivos lineales, que pueden ser recuperados a partir de una secuencia de bits codificada, sin decodificar completamente la secuencia de bits para producir audio, pueden beneficiarse también de los aspectos de la presente invención.

Error en el cálculo de la potencia a partir de los exponentes de Dolby Digital

15 Los exponentes de Dolby Digital $E(k)$ representan una cuantificación tosca del logaritmo de los coeficientes espectrales de TCMD. Hay un número de fuentes de error cuando se usan estos valores como un espectro tosco de potencia.

En primer lugar, en Dolby Digital, el propio procedimiento de cuantificación resulta en un error medio de aproximadamente 2,7 dB, cuando se comparan los valores del espectro de potencia generado a partir de los exponentes (véase la Ecuación 1 anterior) y los valores de potencia calculados directamente a partir de los coeficientes TCMD. Este error medio, que ha sido determinado experimentalmente, puede ser incorporado al desplazamiento constante C en la Ecuación 7 anterior.

20 En segundo lugar, bajo ciertas condiciones de la señal, tales como transitorios, los valores de los exponentes se agrupan a lo largo de la frecuencia (denominados modos "D25" y "D45" en el documento A/52A indicado anteriormente). Esta agrupación a lo largo de la frecuencia hace que el error medio de exponente sea menos predecible y, de esta manera, más difícil de tener en cuenta mediante su incorporación en la constante C de la Ecuación 7. En la práctica, el error debido a este agrupamiento puede ser ignorado por dos razones: (1) la agrupación se usa en pocas ocasiones y (2) la naturaleza de las señales para las que se usa la agrupación resulta en un error medio de medición que es similar al del caso no promediado.

Implementación

30 La invención puede ser implementada en hardware o software, o una combinación de ambos (por ejemplo, matrices, lógicas programables). Si no se especifica lo contrario, los algoritmos o procedimientos incluidos como parte de la invención no están intrínsecamente relacionados a ningún ordenador u otros aparatos. En particular, varias máquinas de propósito general pueden ser usadas con programas escritos según las enseñanzas de la presente memoria, o puede ser más conveniente construir aparatos más especializados (por ejemplo, circuitos integrados) para realizar las etapas requeridas del procedimiento. De esta manera, la invención puede ser implementada en uno o más programas de ordenador que se ejecutan en uno o más sistemas de ordenadores programables, comprendiendo cada uno al menos un procesador, al menos un sistema de almacenamiento de datos (incluyendo memoria volátil y no volátil y/o elementos de almacenamiento), al menos un dispositivo de entrada o puerto y al menos un dispositivo de salida o puerto. El código del programa es aplicado a los datos de entrada para realizar las funciones descritas en la presente memoria y generar la información de salida. La información de salida es aplicada a uno o más dispositivos de salida, en una manera conocida.

45 Cada uno de estos programas puede ser implementado en cualquier lenguaje de programación deseado (incluyendo los lenguajes máquina, ensamblador (assembler) o procedimientos de alto nivel, lógica o de programación orientada a objetos) para comunicarse con un sistema informático. En cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado.

50 Se apreciará que algunas etapas o funciones mostradas en las figuras ejemplares realizan múltiples subetapas y pueden ser mostradas también como múltiples etapas o funciones, en lugar de una etapa o una función. También se apreciará que los diversos dispositivos, funciones, etapas y procedimientos mostrados y descritos en varios ejemplos en la presente memoria pueden ser mostrados combinados o separados, de maneras diferentes a la mostrada en las diversas figuras. Por ejemplo, cuando se implementan por medio de secuencias de instrucciones de software informático, varias funciones y etapas de las figuras ejemplares pueden ser implementadas por medio de secuencias de instrucciones de software multihilo ejecutándose en un hardware de procesamiento de señales digitales adecuado, en cuyo caso los diversos dispositivos y funciones en los ejemplos mostrados en las figuras pueden corresponder a porciones de las instrucciones de software.

60 Preferentemente, cada uno de dichos programas de ordenador es almacenado en, o es descargado a, un dispositivo o un medio de almacenamiento (por ejemplo, medio o memoria de estado sólido, o medio magnético u óptico) que puede ser leído por un ordenador programable de propósito general o especial, para configurar y operar el ordenador cuando el dispositivo o medio de almacenamiento es leído por el sistema informático para realizar los procedimientos descritos en la presente memoria. Puede considerarse también que el sistema de la invención como implementado como un medio de

almacenamiento que puede ser leído por ordenador, configurado con un programa de ordenador, en el que el medio de almacenamiento, configurado de esta manera, hace que un sistema informático opere en una manera específica y predefinida para realizar las funciones descritas en la presente memoria.

- 5 Se han descrito una serie de realizaciones de la invención. No obstante, se entenderá que pueden realizarse diversas modificaciones dentro del alcance de las reivindicaciones adjuntas. Por ejemplo, algunas de las etapas descritas en la presente memoria pueden ser independientes del orden y, de esta manera, pueden ser realizadas en un orden diferente al descrito.

REIVINDICACIONES

- 5 1. Procedimiento para medir la intensidad de una señal de audio codificada en una secuencia de bits, que incluye datos a partir de los cuales puede derivarse una aproximación de un espectro de potencia de la señal de audio, sin decodificar completamente la señal de audio, incluyendo dichos datos representaciones toscas de la señal de audio y representaciones más finas de la señal de audio, siendo seleccionadas dichas representaciones toscas de entre un grupo que contiene factores de escala, envolventes espectrales y coeficientes predictivos lineales, comprendiendo el procedimiento derivar dicha aproximación del espectro de potencia de la señal de audio a partir de dichas representaciones toscas de la señal de audio en dicha secuencia de bits, sin decodificar completamente la señal de audio, y
- 10 determinar una intensidad aproximada de la señal de audio en respuesta a la aproximación del espectro de potencia de la señal de audio.
- 15 2. Procedimiento según la reivindicación 1, en el que la señal de audio codificada en una secuencia de bits es una señal de audio codificada en sub-bandas que tiene una pluralidad de sub-bandas de frecuencia, teniendo cada sub-banda un factor de escala y datos de muestra asociados con el mismo, y en el que las representaciones toscas de la señal de audio comprenden factores de escala y las representaciones más finas asociadas de la señal de audio comprenden datos de muestra asociados con cada factor de escala.
- 20 3. Procedimiento según la reivindicación 2, en el que el factor de escala y los datos de muestra de cada sub-banda representan coeficientes espectrales en la sub-banda, por medio de una notación exponencial en la que el factor de escala comprende un exponente y los datos de muestra asociados comprenden mantisas.
- 25 4. Procedimiento según cualquiera de las reivindicaciones 1-3, en el que dicha secuencia de bits es una secuencia de bits con codificación AC-3.
- 30 5. Procedimiento según la reivindicación 1, en el que la señal de audio codificada en una secuencia de bits es una señal de audio con codificación predictiva lineal, en el que las representaciones toscas de la señal de audio comprenden coeficientes predictivos lineales y las representaciones más finas de la señal de audio comprenden información de excitación asociada con los coeficientes predictivos lineales.
- 35 6. Procedimiento según la reivindicación 1, en el que las representaciones toscas de la señal de audio comprenden al menos una envolvente espectral y las representaciones más finas de la señal de audio comprenden componentes espectrales asociados con dicha al menos una envolvente espectral.
- 40 7. Procedimiento según cualquiera de las reivindicaciones 1-6, en el que la determinación de una intensidad aproximada de la señal de audio en respuesta a la aproximación del espectro de frecuencia de la señal de audio incluye aplicar una medición de intensidad de potencia ponderada.
- 45 8. Procedimiento según la reivindicación 7, en el que la medición de intensidad de potencia ponderada emplea un filtro que resta importancia a las frecuencias menos perceptibles y realiza un promedio de la potencia de la señal de audio filtrada en el tiempo.
- 50 9. Procedimiento según cualquiera de las reivindicaciones 1-6, en el que la determinación de una intensidad aproximada de la señal de audio en respuesta a la aproximación del espectro de frecuencia de la señal de audio incluye aplicar una medición de intensidad psicoacústica.
- 55 10. Procedimiento según la reivindicación 9, en el que la medición de intensidad psicoacústica emplea un modelo del oído humano para determinar una intensidad específica en cada una de entre una pluralidad de bandas de frecuencia similares a las bandas críticas del oído humano.
- 60 11. Procedimiento según la reivindicación 9 y una cualquiera de las reivindicaciones 2 y 3, en el que dichas sub-bandas son similares a las bandas críticas del oído humano y la medición de intensidad psicoacústica emplea un modelo del oído humano para determinar una intensidad específica en cada una de dichas sub-bandas.
12. Aparato para medir la intensidad de una señal de audio codificada en una secuencia de bits que incluye datos a partir de los cuales puede derivarse una aproximación de un espectro de frecuencia de la señal de audio, sin decodificar completamente la señal de audio, incluyendo dichos datos representaciones toscas de la señal de audio y representaciones más finas asociadas de la señal de audio, siendo seleccionadas dichas representaciones toscas de entre un grupo que contiene factores de escala, envolventes espectrales y coeficientes predictivos lineales, comprendiendo el aparato

- medios (502) para derivar dicha aproximación del espectro de potencia de la señal de audio a partir de las representaciones toscas de la señal de audio en dicha secuencia de bits, sin decodificar completamente la señal de audio, y
 5 medios (504) para determinar una intensidad aproximada de la señal de audio en respuesta a la aproximación del espectro de potencia de la señal de audio.
13. Aparato según la reivindicación 12, en el que la señal de audio codificada en una secuencia de bits es una señal de audio codificada en sub-bandas, que tiene una pluralidad de sub-bandas de frecuencia, teniendo cada sub-banda un factor de escala y datos de muestra asociados con el mismo, y en el que las representaciones toscas de la señal de audio comprenden factores de escala y las representaciones más finas asociadas de la señal de audio comprenden datos de muestra asociados con cada factor de escala.
 10
14. Aparato según la reivindicación 13, en el que el factor de escala y los datos de muestra de cada sub-banda representan coeficientes espectrales en la sub-banda por medio de una notación exponencial en la que el factor de escala comprende un exponente y los datos de muestra asociados comprenden mantisas.
 15
15. Aparato según cualquiera de las reivindicaciones 12-14, en el que dicha secuencia de bits es una secuencia de bits con codificación AC-3.
16. Aparato según la reivindicación 12, en el que la señal de audio codificada en una secuencia de bits es una señal de audio con codificación predictiva lineal, en el que las representaciones toscas de la señal de audio comprenden coeficientes predictivos lineales y las representaciones más finas de la señal de audio comprenden información de excitación asociada con los coeficientes predictivos lineales.
 20
17. Aparato según la reivindicación 12, en el que las representaciones toscas de la señal de audio comprenden al menos una envolvente espectral y las representaciones más finas de la señal de audio comprenden componentes espectrales asociados con dicha al menos una envolvente espectral.
 25
18. Aparato según cualquiera de las reivindicaciones 12-17, en el que dichos medios para determinar una intensidad aproximada de la señal de audio, en respuesta a la aproximación del espectro de frecuencia de la señal de audio, incluyen medios (601) para aplicar una medición de intensidad de potencia ponderada.
 30
19. Aparato según la reivindicación 18, en el que la medición de intensidad de potencia ponderada emplea un filtro que resta importancia a las frecuencias menos perceptibles y promedia la potencia de la señal de audio filtrada en el tiempo.
 35
20. Aparato según cualquiera de las reivindicaciones 12-17, en el que dichos medios (504) para determinar una intensidad aproximada de la señal de audio, en respuesta a la aproximación del espectro de frecuencia de la señal de audio, incluyen medios para aplicar una medición de intensidad psicoacústica.
21. Aparato según la reivindicación 20, en el que la medición de intensidad psicoacústica emplea un modelo del oído humano para determinar una intensidad específica en cada una de entre una pluralidad de bandas de frecuencia similares a las bandas críticas del oído humano.
 40
22. Aparato según la reivindicación 20 y una cualquiera de las reivindicaciones 13 y 14, en el que dichas sub-bandas son similares a las bandas críticas del oído humano y la medición de intensidad psicoacústica emplea un modelo del oído humano para determinar una intensidad específica en cada una de dichas sub-bandas.
 45
23. Aparato adaptado para realizar los procedimientos de una cualquiera de las reivindicaciones 1 a 11.
24. Programa de ordenador, almacenado en un medio que puede ser leído por ordenador, para hacer que un ordenador realice el procedimiento según una cualquiera de las reivindicaciones 1 a 11.
 50

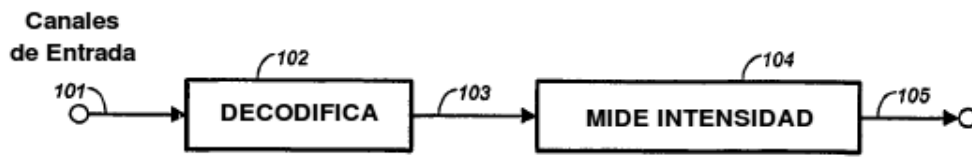


FIG. 1

TÉCNICA ANTERIOR

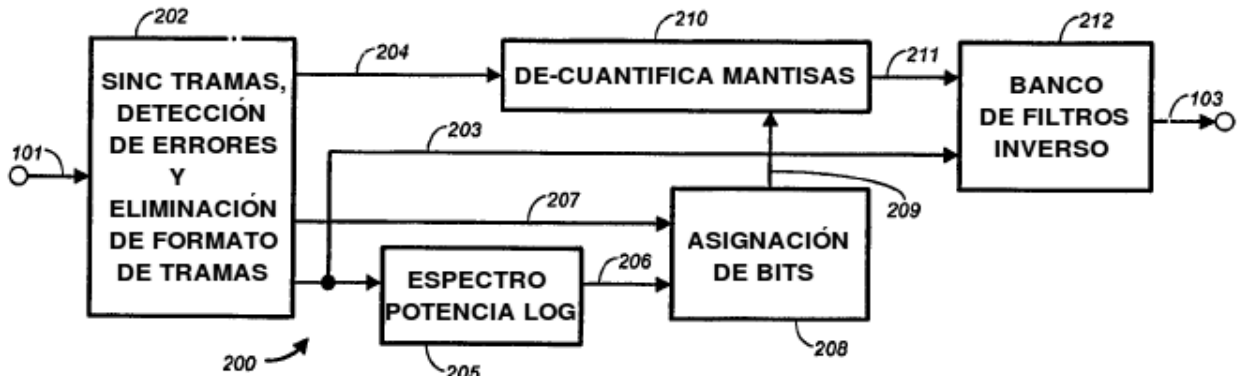


FIG. 2

TÉCNICA ANTERIOR

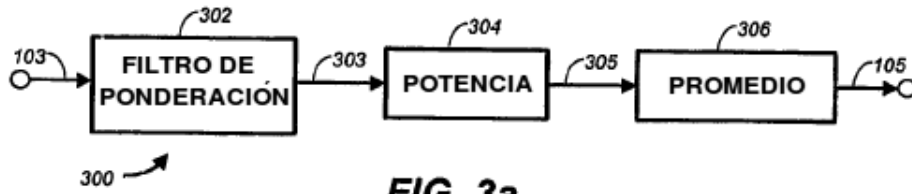


FIG. 3a

TÉCNICA ANTERIOR

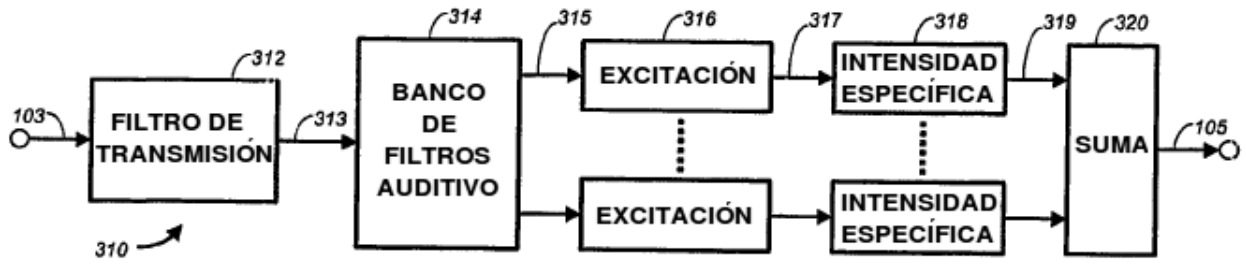


FIG. 3b

TÉCNICA ANTERIOR

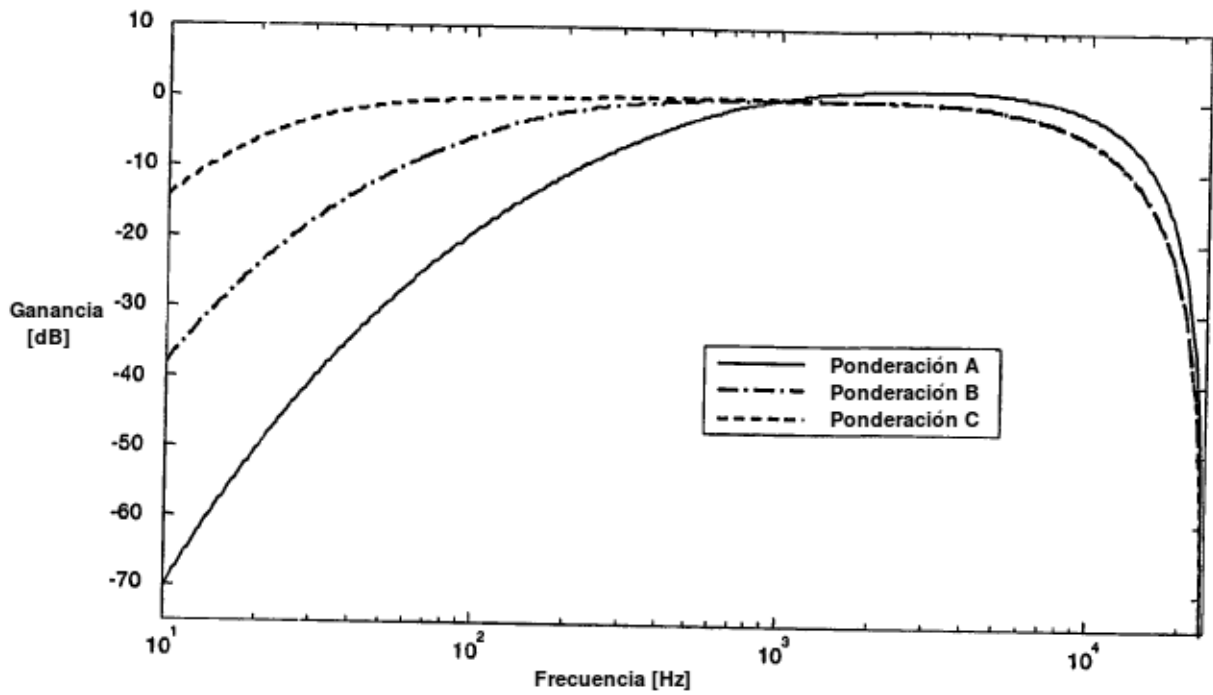


FIG. 4
TÉCNICA ANTERIOR

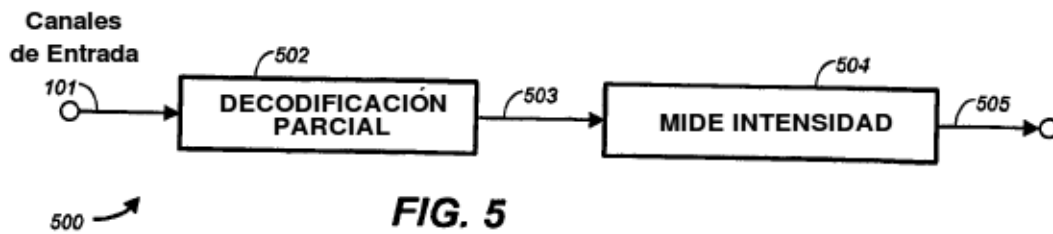


FIG. 5

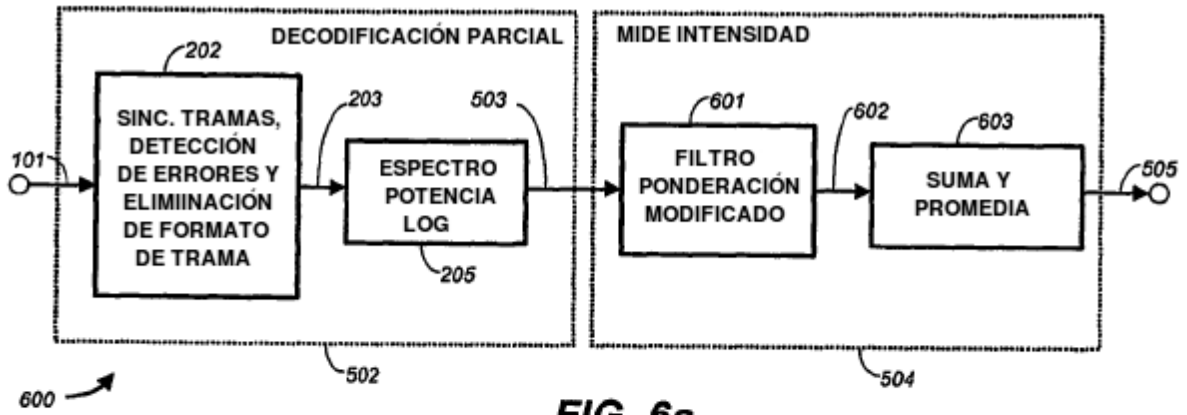


FIG. 6a

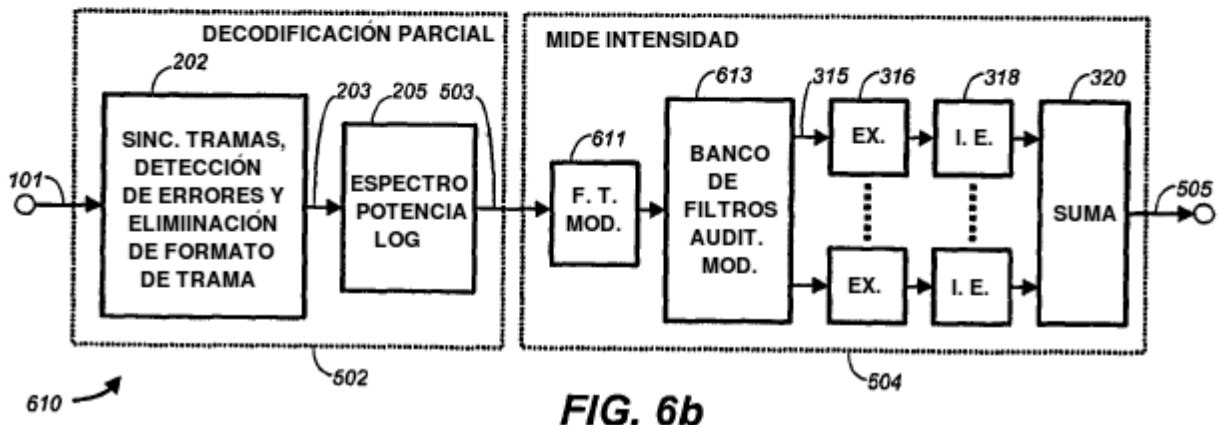


FIG. 6b