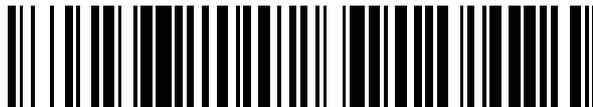


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 526 761**

51 Int. Cl.:

G10L 25/69 (2013.01)

G10L 21/0264 (2013.01)

H03G 9/00 (2006.01)

H03G 9/02 (2006.01)

G10L 21/02 (2013.01)

G10L 21/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.04.2010 E 10160679 (6)**

97 Fecha y número de publicación de la concesión europea: **03.12.2014 EP 2381574**

54 Título: **Aparato y método para modificar una señal de audio de entrada**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.01.2015

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27C
80686 München, DE**

72 Inventor/es:

**UHLE, CHRISTIAN;
HERRE, JUERGEN;
HELLMUTH, OLIVER y
FINAUER, STEFAN**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 526 761 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

Aparato y método para modificar una señal de audio de entrada

5 **[0001]** Las realizaciones de acuerdo con la invención se refieren al procesamiento de una señal de audio y, particularmente, a un aparato y método para modificar una señal de audio de entrada.

10 **[0002]** Ha habido muchos intentos de desarrollar un método objetivo satisfactorio para medir el volumen. Fletcher y Munson determinaron en 1933 que el oído humano es menos sensible frecuencias altas y bajas que a frecuencias medias (o voz). También encontraron que el cambio relativo en la sensibilidad disminuía conforme el nivel del sonido se incrementaba. Un medidor del volumen inicial consistía de un micrófono, amplificador, medidor y una combinación de filtros diseñados para imitar aproximadamente la respuesta a la frecuencia del oído a niveles de sonido bajos, medios y altos.

15 **[0003]** Aunque tales dispositivos proporcionaron una medición del volumen de un tono aislado de un solo nivel constante, las mediciones de sonidos más complejos no correspondían muy bien a las impresiones subjetivas del volumen. Los medidores del nivel del sonido de este tipo se han estandarizado, pero sólo se utilizan para tareas específicas, tales como la verificación y control del ruido industrial.

20 **[0004]** A inicios de la década de 1950, Zwicker y Stevens, entre otros, extendieron el trabajo de Fletcher y Munson al desarrollar un modelo más realista del proceso de percepción del volumen. Stevens publicó un método para el "Cálculo del Volumen del Ruido Complejo" en la revista de la Sociedad Acústica de América en 1956, y Zwicker publicó su artículo "Base Sicológica y Metódica del Volumen" en Acoustica en 1958. En 1959, Zwicker publicó un procedimiento gráfico para el cálculo del volumen, así como varios artículos similares poco después. Los métodos de Stevens y Zwicker se estandarizaron como ISO 532, partes A y B (respectivamente). Ambos métodos involucran pasos similares.

30 **[0005]** Primero, la distribución que varía con el tiempo de la energía a lo largo de una membrana basilar del oído interno, referida como excitación, se simuló pasando el audio a través de un banco de filtros auditivos de paso de banda, con frecuencias centrales separadas de manera uniforme en una escala de relación de la banda crítica. Cada filtro auditivo se diseñó para simular la respuesta a la frecuencia en una ubicación particular a lo largo de la membrana basilar del oído interno, con la frecuencia central del filtro que corresponde a esta ubicación. Un ancho de banda crítica se define como el ancho de banda de tal filtro. Medida en unidades de Hertz, el ancho de banda crítica de estos filtros auditivos se incrementa con la frecuencia central que se incrementa. Por lo tanto, es útil definir una escala de frecuencia deformada, de manera que el ancho de banda crítica para todos los filtros auditivos medidos en esta escala deformada sea constante. Tal escala deformada se refiere como la escala de relación de la banda crítica, y es muy útil para entender y simular una gama de fenómenos psicoacústicos. Véase, por ejemplo, Psychoacoustics-Facts and Models por E. Zwicker y H. Fastl, Springer-Verlag, Berlín, 1990. Los métodos de Stevens y Zwicker utilizan una escala de relación de la banda crítica referida como la escala Bark, en la cual el ancho de banda crítica es constante por debajo de 500 Hz, y se incrementa por encima de 500 Hz. Más recientemente, Moore y Glasberg definieron una escala de relación de la banda crítica, que nombraron la escala del Ancho de Banda Rectangular Equivalente (ERB) (B. C. J. Moore, B. Glasberg, T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness," Journal of the Audio Engineering Society, Vol. 45, No. 4, Abril 1997, pp. 224-240). A través de experimentos psicoacústicos que utilizan enmascaradores del ruido con muescas, Moore y Glasberg demostraron que el ancho de banda crítica continúa disminuyendo por debajo de 500 Hz, en contraste con la escala Bark, en donde el ancho de banda crítica permanece constante.

50 **[0006]** El término "banda crítica" se remonta al trabajo de Harvey Fletcher en 1938, sobre el enmascaramiento de la sensación del sonido por señales acompañantes ("J. B. Allen, "A short history of telephone psychophysics", Audio Eng. Soc. Convention, 1997"). Las bandas críticas pueden expresarse utilizando la escala Bark propuesta por Zwicker en 1961: cada banda crítica tiene el ancho de un Bark (una unidad nombrada por Heinrich Barkhausen). Sobre los bancos de filtros que imitan la percepción auditiva humana, existe, por ejemplo, la escala del Ancho de Banda Rectangular Equivalente (ERB) ("B. C. J. Moore, B. R. Glasberg and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness", J. Audio Eng. Soc., 1997").

55 **[0007]** El término "volumen específico", describe la sensación del volumen causado por una señal en una cierta región de la membrana basilar a un cierto ancho de banda de la frecuencia medido en las bandas críticas. Se mide en unidades de Sone/Bark. El término "banda crítica", se relaciona con las bandas de frecuencia de un banco de filtros auditivos, que comprende bancos de filtros de paso de banda no uniforme, diseñados para imitar la resolución de la frecuencia del oído humano. El volumen total de un sonido equivale a la suma/integral del volumen específico a través de todas las bandas críticas.

60 **[0008]** Un método para procesar una señal de audio ha sido descrito en "A. J. Seefeldt, "Cálculo y ajuste del volumen percibido y/o el equilibrio espectral percibido de una señal de audio" (Calculating and adjusting the perceived loudness and/or the perceived spectral balance of an audio signal). Patente de los Estados Unidos 2009/0097676, 2009". Este método tiene el objeto de controlar el volumen específico de la señal de audio, con

aplicaciones para controlar el volumen, controlar el intervalo dinámico, igualación dinámica y compensación del ruido de fondo. En este documento, una señal de audio de entrada (normalmente en el dominio de la frecuencia), se modifica de manera que su volumen específico corresponde el volumen específico objetivo.

5 **[0009]** Para ilustrar la ventaja del procesamiento presentado en “A. J. Seefeldt, “Cálculo y ajuste del volumen percibido y/o el equilibrio espectral percibido de una señal de audio” (Calculating and adjusting the perceived loudness and/or the perceived spectral balance of an audio signal). Patente de los Estados Unidos 2009/0097676, 2009”, considerar el control del volumen de una señal de audio. El cambiar el nivel de una señal de audio en la reproducción del sonido normalmente está dirigido al cambio de su volumen percibido. Dicho de manera diferente, el control del volumen es implementado tradicionalmente como el control del nivel del sonido. Sin embargo, nuestra experiencia diaria y el conocimiento de la psicoacústica, indican que esto no óptimo.

15 **[0010]** La sensibilidad del oído humano varía con la frecuencia y el nivel, de manera que una disminución del nivel de la intensidad del sonido atenúa la sensación de las frecuencias bajas y altas (por ejemplo, alrededor de 100 Hz y 10000 Hz, respectivamente), más que la sensación de las frecuencias medias (por ejemplo, entre 2000 y 4000 Hz). Cuando se disminuye el nivel de reproducción de un nivel “cómodamente fuerte” (por ejemplo, 75-80 dBA) a un nivel más bajo, por ejemplo, 18 dB, el equilibrio espectral percibido de la señal de audio cambia. Esto se ilustra en los bien conocidos Contornos de Volumen Igual, referidos con frecuencia como las Curvas de Fletcher-Munson (por los investigadores que midieron primero los Contornos del Volumen Igual en 1933). El Contorno de Volumen Igual muestra el nivel de presión del sonido (SPL) sobre el espectro de la frecuencia, para el cual un oyente percibe un volumen constante cuando se presenta con tonos constantes puros.

25 **[0011]** Los Contornos de Volumen Igual se describen en, por ejemplo “B. C. J. Moore, B. R. Glasberg y T. Baer, “A model for the prediction of thresholds, loudness, and partial loudness”, J. Audio Eng. Soc., 1997), p. 232, Figure 13”. Una medición revisada se ha estandarizado como ISO 226:2003 en el 2003.

30 **[0012]** En consecuencia, el control del volumen convencional no sólo cambia el volumen sino también el timbre. El impacto de este efecto depende del SPL (es menos pronunciado cuando se cambia el SPL de, por ejemplo, 86 dBA a 68 dBA, en comparación con un cambio de 76 dBA a 58 dBA), pero no es deseado en todas las clases.

[0013] Esto se compensa por el procesamiento descrito en “A. J. Seefeldt, “Cálculo y ajuste del volumen percibido y/o el equilibrio espectral percibido de una señal de audio” (Calculating and adjusting the perceived loudness and/or the perceived spectral balance of an audio signal). Patente de los Estados Unidos 2009/0097676, 2009”.

35 **[0014]** La Figura 7 muestra un diagrama de flujo de un método 700 descrito en “A. J. Seefeldt, “Cálculo y ajuste del volumen percibido y/o el equilibrio espectral percibido de una señal de audio” (Calculating and adjusting the perceived loudness and/or the perceived spectral balance of an audio signal). Patente de los Estados Unidos 2009/0097676, 2009”.

40 **[0015]** La señal de salida se procesa calculando en 710 la señal de excitación, calculando en 720 el volumen específico, calculando en 730 el volumen específico objetivo, calculando en 740 la señal de excitación objetivo, calculando en 750 las ponderaciones espectrales y aplicando en 760 las ponderaciones espectrales a la señal de entrada y resintetizando la señal de salida.

45 **[0016]** Las ponderaciones espectrales H son ponderaciones de las bandas de frecuencia que dependen del volumen específico de la señal de entrada y del volumen específico objetivo. Su cálculo, descrito en “A. J. Seefeldt, “Cálculo y ajuste del volumen percibido y/o el equilibrio espectral percibido de una señal de audio” (Calculating and adjusting the perceived loudness and/or the perceived spectral balance of an audio signal). Patente de los Estados Unidos 2009/0097676, 2009”, comprende el cálculo del volumen específico y el proceso inverso del cálculo del volumen específico, que se aplica al volumen específico objetivo.

50 **[0017]** Ambos pasos de procesamiento introducen una alta carga computacional. Los métodos para el cálculo del volumen específico se han presentado en “E. Zwicker, H. Fastl, U. Widmann, K. Kurakata, S. Kuwano y S. Namba, “Program for calculating loudness according to DIN 45631 (ISO 532 B)”, J. Acoust. Soc. Jpn. (E), vol. 12, 1991” y “B. C. J. Moore, B. R. Glasberg y T. Baer, “A model for the prediction of thresholds, loudness, and partial loudness”, J. Audio Eng. Soc., 1997”.

55 **[0018]** Es el objeto de la presente invención proporcionar un concepto mejorado para modificar las señales de audio, para permitir una implementación eficiente con una baja complejidad computacional.

60 **[0019]** Este objeto se soluciona por un aparato de acuerdo con la reivindicación 1 o un método de acuerdo con la reivindicación 20.

65 **[0020]** Una realización de la invención proporciona un aparato para modificar una señal de audio de entrada, que comprende un determinador de la excitación, un dispositivo de almacenamiento y un modificador de la señal. El determinador de la excitación está configurado para determinar un valor de un parámetro de la excitación de una

subbanda de una pluralidad de subbandas de la señal de audio de entrada, basándose en el contenido de energía de la subbanda, en donde el valor del parámetro de la excitación indica una potencia, una energía de corto plazo, o un valor cuantificado de la energía de corto plazo de la señal de audio en la subbanda. El dispositivo de almacenamiento está configurado para almacenar una tabla de búsqueda que contiene una pluralidad de factores de ponderación espectral, en donde un factor de ponderación espectral de la pluralidad de los factores de ponderación espectral se asocia con un valor predefinido del parámetro de la excitación y con una subbanda de la pluralidad de subbandas. Además, el dispositivo de almacenamiento está configurado para proporcionar un factor de ponderación espectral que corresponde al valor determinado del parámetro de la excitación y que corresponde a la subbanda, para la cual se determina el valor del parámetro de la excitación. El modificador de la señal está configurado para modificar un contenido de la subbanda de la señal de audio de entrada, para la cual se determina el parámetro de la excitación, basándose en el factor de ponderación espectral proporcionado para proporcionar una subbanda modificada mediante el escalado multiplicativo de la subbanda de la señal de audio por el factor de ponderación espectral proporcionado por la tabla de búsqueda.

[0021] Las realizaciones de acuerdo con la presente invención se basan en la idea central de que las subbandas de una señal de audio pueden modificarse fácilmente, utilizando una tabla de búsqueda que contiene factores de ponderación espectral, que pueden elegirse dependiendo de la subbanda respectiva y del parámetro de la excitación de la subbanda. Para esto, la tabla de búsqueda contiene factores de ponderación espectral para una pluralidad de valores predefinidos del parámetro de la excitación, para al menos una subbanda predefinida de la pluralidad de subbandas. Al utilizar la tabla de búsqueda, la complejidad computacional puede reducirse de manera significativa, puesto que un cálculo explícito de los factores de ponderación espectral (que incluye el cálculo del volumen, su modificación y el procesamiento inverso del cálculo del volumen) no es necesario. Por lo tanto, se habilita una implementación eficiente.

[0022] En algunas realizaciones de acuerdo con la invención, el determinador de la excitación determina un valor de un parámetro de la excitación, no para todas las subbandas de la pluralidad de subbandas. Además, la tabla de búsqueda contiene sólo factores de ponderación espectral asociados con las subbandas, para las cuales se determina un valor del parámetro de la excitación. De esta manera, el espacio de almacenamiento requerido de la tabla de búsqueda y el esfuerzo computacional para el determinador de la excitación pueden reducirse.

[0023] Algunas realizaciones de acuerdo con la invención se relacionan con una tabla de búsqueda que comprende exactamente tres dimensiones, asociadas con valores predefinidos del parámetro de la excitación, con subbandas de la pluralidad de subbandas y con valores predefinidos de un parámetro de modificación externa.

[0024] Algunas realizaciones adicionales de acuerdo con la invención, se relacionan con una tabla de búsqueda que comprende exactamente cuatro dimensiones, asociadas con valores predefinidos del parámetro de la excitación, con subbandas de la pluralidad de subbandas, con valores predefinidos del parámetro de modificación externa y con valores predefinidos de un parámetro del ruido de fondo.

[0025] Las realizaciones de acuerdo con la invención se detallarán posteriormente, refiriéndose a los dibujos anexos, en los cuales:

La Figura 1 es un diagrama de bloques de un aparato para modificar una señal de audio de entrada;

La Figura 2 es una ilustración esquemática de los contornos del volumen igual;

La Figura 3 es una ilustración esquemática de los contornos del volumen igual, normalizados por filtros de transmisión;

La Figura 4 es un diagrama de bloques de un aparato para modificar una señal de audio de entrada;

La Figura 5 es un diagrama de flujo de un método para modificar una señal de audio de entrada;

La Figura 6 es un diagrama de flujo de un método para modificar una señal de audio de entrada; y

La Figura 7 es un diagrama de flujo de un método conocido para modificar una señal de audio de entrada.

[0026] En lo siguiente, los mismos números de referencia se utilizan parcialmente para los objetos y unidades funcionales que tienen las mismas o similares propiedades funcionales y la descripción de los mismos con respecto a una Figura, se aplicará también a otras Figuras, con el fin de reducir la redundancia en la descripción de las realizaciones.

[0027] La Figura 1 muestra un diagrama de bloques de un aparato 100 para modificar una señal de la subbanda de audio de entrada 102 de acuerdo con una realización de la invención. El aparato 100 comprende un determinador de la excitación 110, un dispositivo de almacenamiento 120 y un modificador de la señal 130. El determinador de la excitación 110 está conectado al dispositivo de almacenamiento 120 y el dispositivo de almacenamiento 120 está

conectado al modificador de la señal 130. El determinador de la excitación 110 determina un valor 112 de un parámetro de la excitación de una subbanda 102 de una pluralidad de subbandas de la señal de audio de entrada 102, basándose en el contenido de energía de la subbanda 102. El dispositivo de almacenamiento 120 almacena una tabla de búsqueda que contiene una pluralidad de factores de ponderación espectral, en donde un factor de ponderación espectral 124 de la pluralidad de factores de ponderación espectral está asociado con un valor predefinido del parámetro de la excitación y con una subbanda de la pluralidad de subbandas. Además, el dispositivo de almacenamiento 120 proporciona un factor de ponderación espectral 124 que corresponde al valor determinado 112 del parámetro de la excitación y que corresponde a la subbanda 102, para la cual se determina el valor 112 del parámetro de la excitación. El modificador de la señal 130 modifica un contenido de la subbanda 102 de la señal de audio de entrada, para la cual se determina el valor 112 del parámetro de la excitación, basándose en el factor de ponderación espectral proporcionado 124, para obtener y proporcionar una subbanda modificada 132.

[0028] Al utilizar una tabla de búsqueda para proporcionar los factores de ponderación espectral 124 para modificar la señal de audio de entrada, la complejidad computacional puede reducirse de manera significativa, en comparación con los conceptos conocidos.

[0029] El determinador de la excitación 110 determina un valor 112 de un parámetro de la excitación basándose en el contenido de energía de la subbanda 102. Esto puede hacerse, por ejemplo, midiendo el contenido de energía de una subbanda 102 para determinar el valor 112 del parámetro de la excitación para la subbanda 102. De esta manera, un parámetro de la excitación puede representar una medida para la potencia por subbanda o una energía de corto plazo en una subbanda específica, puesto que el contenido de energía puede variar con el tiempo y/o entre diferentes subbandas. De manera alterna, el valor del parámetro de la excitación puede determinarse basándose en una función (única, inyectiva, biyectiva) de la energía de corto plazo de una subbanda (por ejemplo, una función exponencial, una función logarítmica o una función lineal). Por ejemplo, puede utilizarse una función de cuantificación. En este ejemplo, el determinador de la excitación 110 puede medir un contenido de energía de la subbanda, y puede cuantificar el contenido de energía medido de la subbanda, para obtener el valor del parámetro de la excitación, de manera que el valor del parámetro de la excitación es igual a un valor predefinido del parámetro de la excitación. En otras palabras, un valor de la energía medida puede asignarse a un valor predefinido del parámetro de la excitación (por ejemplo, el valor predefinido más cercano del parámetro de la excitación). De manera alterna, el valor del parámetro de la excitación indica directamente el contenido de energía medido y el dispositivo de almacenamiento 120 puede asignar el valor determinado del parámetro de la excitación a un valor predefinido del parámetro de la excitación.

[0030] Las subbandas de la señal de audio de entrada pueden representar diferentes bandas de frecuencia de la señal de audio de entrada. Para tomar en cuenta una distribución perceptual de las bandas de frecuencia, las subbandas pueden distribuirse, por ejemplo, de acuerdo con la escala ERB o la escala Bark u otra separación frecuencial que imite la resolución frecuencial del oído humano. En otras palabras, las subbandas de la pluralidad de subbandas de la señal de audio de entrada pueden dividirse de acuerdo con la escala ERB o la escala Bark.

[0031] El dispositivo de almacenamiento 120 comprende una entrada para el parámetro de la excitación (señal de excitación) y para un índice de la subbanda que indica la subbanda 102, para la cual se determina el valor 112 del parámetro de la excitación. De manera alterna, el dispositivo de almacenamiento comprende una o más entradas adicionales para los parámetros adicionales.

[0032] El dispositivo de almacenamiento 120 puede ser un medio de almacenamiento digital, como por ejemplo, una memoria de sólo lectura (ROM), un disco duro, un CD, un DVD o cualquier otra clase de memoria no volátil, o una memoria de acceso aleatorio (RAM).

[0033] La tabla de búsqueda representa al menos una matriz bidimensional que contiene la pluralidad de factores de ponderación espectral. Un factor de ponderación espectral 124 contenido en la tabla de búsqueda, está asociado de manera no ambigua a un valor predefinido del parámetro de la excitación y a una subbanda de la pluralidad de subbandas. En otras palabras, cada factor de ponderación espectral contenido en la tabla de búsqueda, puede asociarse a un valor predefinido del parámetro de la excitación y una subbanda de la pluralidad de subbandas. El dispositivo de almacenamiento 120 puede proporcionar un factor de ponderación espectral 124 asociado con un valor predefinido del parámetro de la excitación más cercano al valor determinado 112 del parámetro de la excitación. De manera alterna, por ejemplo, el dispositivo de almacenamiento 120 puede interpolar de manera lineal o logarítmica los dos factores de ponderación espectral asociados con los dos valores predefinidos del parámetro de la excitación más cercano al valor determinado 112 del parámetro de la excitación.

[0034] Los valores predefinidos del parámetro de la excitación pueden distribuirse de manera lineal o logarítmica.

[0035] El modificador de la señal 130 puede, por ejemplo, amplificar o atenuar el contenido de la subbanda 102, para la cual se determina el valor 112 del parámetro de la excitación, por el factor de ponderación espectral proporcionado 124.

[0036] Al utilizar el concepto descrito, por ejemplo, una atenuación variable de la atenuación del oído humano de las frecuencias bajas, medias y altas causada por un incremento o del nivel de intensidad del sonido de una señal de audio, puede compensarse fácilmente. Por ejemplo, cuando se disminuye el nivel de reproducción de un nivel a otro nivel, el equilibrio espectral percibido de la señal de audio cambia. Esto se ilustra en la Figura 2 y la Figura 3, que representan los contornos del volumen igual. Especialmente en la región de baja frecuencia, los contornos de diferentes volúmenes iguales no son paralelos unos con otros. Una amplificación o atenuación de la banda de frecuencia baja diferente de las bandas de frecuencia media y/o alta, puede ser igual a una flexión de los contornos del volumen igual, de manera que pueden ser paralelos o más paralelos que antes. De esta manera, el cambio del equilibrio espectral percibido puede compensarse o casi compensarse utilizando el concepto descrito.

[0037] La diferencia entre los contornos del volumen igual de la Figura 2 y los contornos del volumen igual de la Figura 3, es una normalización mediante un filtro de transmisión. Este filtro de transmisión puede simular un efecto de filtración de la transmisión del audio a través del oído externo e interno. Tal filtro de transmisión puede implementarse opcionalmente en un aparato mostrado en la Figura 1 para filtrar la señal de audio de entrada antes de proporcionarla al determinador de la excitación 110.

[0038] Para una modificación más continua de la señal de audio de entrada, el determinador de la excitación 110 puede determinar un valor 112 de un parámetro de la excitación para más de una subbanda de la pluralidad de subbandas. A continuación, el dispositivo de almacenamiento 120 puede proporcionar un factor de ponderación espectral 124 para cada subbanda 102, para la cual se determina el valor 112 de un parámetro de la excitación, y el modificador de la señal 130 puede modificar un contenido de cada subbanda 102, para la cual se determina el valor 112 de un parámetro de la excitación, basándose en el factor de ponderación espectral 124 proporcionado correspondiente.

[0039] La pluralidad de las subbandas de la señal de audio de entrada puede proporcionarse mediante una unidad de memoria o puede generarse mediante por un banco de filtros de análisis.

[0040] Un parámetro de la excitación puede determinarse para una subbanda, para más de una subbanda o para todas las subbandas de la pluralidad de subbandas. Para esto, el aparato 100 puede comprender sólo un determinador de la excitación 110 que determina uno, más de uno o todos los valores de los parámetros de la excitación o puede comprender un determinador de la excitación 110 para cada subbanda 102, para la cual se determina el valor 112 de un parámetro de la excitación. Además, el aparato 100 puede comprender uno o más modificadores de la señal 130 para una o más subbandas, para las cuales se determina un parámetro de la excitación. Sin embargo, es suficiente utilizar una sola tabla de búsqueda (y un dispositivo de almacenamiento) para todas las subbandas 102, para las cuales se determina el valor 112 de un parámetro de la excitación.

[0041] El determinador de la excitación 110, el dispositivo de almacenamiento 120 y el modificador de la señal pueden ser unidades de elementos físicos independientes, parte de una computadora, microcontrolador o procesador de la señal digital, así como un programa para computadora o un producto de programas configurado para correrse en una computadora, microcontrolador o procesador de la señal digital.

[0042] La Figura 4 muestra un diagrama de bloques de un aparato 400 para modificar una señal de audio de entrada de acuerdo con una realización de la invención. El aparato 400 es similar al aparato mostrado en la Figura 1, pero comprende además, un banco de filtros de análisis 410 y un banco de filtros de síntesis 420. El banco de filtros de análisis 410 separa la señal de audio de entrada en la pluralidad de subbandas. A continuación, el determinador de la excitación 110, determina un valor del parámetro de la excitación (calcula una característica) para una o más subbandas de la pluralidad de subbandas. Posteriormente, el dispositivo de almacenamiento 120 proporciona uno o más factores de ponderación espectral correspondientes a uno o más modificadores de la señal 130. Finalmente, el banco de filtros de síntesis 420 combina la pluralidad de subbandas que contienen al menos una subbanda modificada, para obtener y proporcionar una señal de audio modificada (o producir una señal de audio).

[0043] El ejemplo mostrado en la Figura 4 puede ser una aplicación del método propuesto para un caso genérico. El procesamiento mostrado para la señal de la n-ésima subbanda (n-ésima subbanda) puede aplicarse a todas las otras señales de la subbanda (o sólo a todas las subbandas, para las cuales se determina un valor del parámetro de la excitación), de la misma manera.

[0044] Opcionalmente, un factor de ponderación espectral contenido en la tabla de búsqueda, está asociado además con un valor predefinido de un parámetro de modificación externa, como se indica por la línea punteada en la Figura 4 (pero también aplicable al aparato mostrado en la Figura 1). El parámetro de modificación externa (o simplemente el parámetro de modificación) puede representar, por ejemplo, un valor de entrada de una interconexión del usuario (por ejemplo, ajustes del volumen y/o ambiente). En consecuencia, en este caso, el dispositivo de almacenamiento 120 puede proporcionar un factor de ponderación espectral que corresponde al valor del parámetro de modificación externa. Por ejemplo, si un usuario incrementa o disminuye el ajuste del volumen, el valor del parámetro de modificación externa cambia y el dispositivo de almacenamiento 120 puede proporcionar otro factor de ponderación espectral correspondiente. Resumiendo, el dispositivo de almacenamiento 120 puede proporcionar un factor de ponderación espectral que corresponde al valor determinado del parámetro de la

excitación de una subbanda, que corresponde a la subbanda, para la cual se determina el valor del parámetro de la excitación, y que corresponde a un valor del parámetro de modificación externa.

5 **[0045]** En este ejemplo, la tabla de búsqueda puede comprender exactamente tres dimensiones asociadas con los valores predefinidos del parámetro de la excitación, asociadas con las subbandas de la pluralidad de subbandas y asociadas con los valores predefinidos del parámetro de modificación externa. Esto significa, que cada factor de ponderación espectral contenido en la tabla de búsqueda, está asociado con un valor predefinido específico del parámetro de la excitación, con una subbanda de la pluralidad de subbandas y con un valor predefinido específico del parámetro de modificación externa. En otras palabras, la tabla de búsqueda contiene para cada combinación de
10 un valor predefinido del parámetro de la excitación, una subbanda y un valor predefinido del parámetro de modificación externa de un factor de ponderación espectral. Los valores predefinidos del parámetro de modificación externa pueden distribuirse, por ejemplo, de manera lineal o logarítmica a través de un posible intervalo del parámetro de modificación externa.

15 **[0046]** Además, en algunas realizaciones, un factor de ponderación espectral contenido en la tabla de búsqueda, está asociado también con un valor predefinido de un parámetro del ruido de fondo. El parámetro del ruido de fondo puede representar el nivel del ruido de fondo de la señal de audio de entrada. De esta manera, por ejemplo, puede realizarse una compensación del efecto del enmascaramiento parcial de una señal de audio en presencia del ruido de fondo. En este caso, el dispositivo de almacenamiento puede proporcionar un factor de ponderación espectral que corresponde a un valor del parámetro del ruido de fondo. Esto puede hacerse de manera adicional o alterna a la consideración mencionada anteriormente del parámetro de modificación externa. Si ambos se consideran, el dispositivo de almacenamiento puede proporcionar el factor de ponderación espectral que corresponde al valor determinado del parámetro de la excitación de la subbanda, que corresponde a la subbanda para la cual se determina el parámetro de la excitación, que corresponde a un valor del parámetro de modificación externa y que
20 corresponde a un valor del parámetro del ruido de fondo. En este caso, la tabla de búsqueda puede comprender exactamente cuatro dimensiones asociadas con valores predefinidos del parámetro de la excitación, asociadas con las subbandas de la pluralidad de subbandas, asociadas con los valores predefinidos del parámetro de modificación externa y asociadas con los valores predefinidos del parámetro del ruido de fondo. Los valores predefinidos del parámetro del ruido de fondo pueden distribuirse, por ejemplo, de manera lineal o logarítmica a través de un posible
25 intervalo del parámetro del ruido de fondo.

[0047] Un valor del parámetro del ruido de fondo puede determinarse mediante un detector del ruido de fondo. Esto puede hacerse para toda la señal de audio de entrada antes de la separación en las subbandas, o a nivel de la subbanda para una subbanda, para más de una subbanda o para todas las subbandas de manera individual. De
35 manera alterna, si la pluralidad de subbandas de la señal de audio de entrada está almacenada y se proporciona por una unidad de memoria, el valor del parámetro del ruido de fondo puede también proporcionarse por la unidad de memoria.

[0048] En cualquier caso, el dispositivo de almacenamiento no comprende una entrada para un parámetro del volumen específico o un parámetro del volumen específico objetivo, aunque los factores de ponderación espectral contenidos en la tabla de búsqueda, pueden calcularse basándose en un parámetro del volumen específico o un parámetro del volumen específico objetivo. El cálculo de los factores de ponderación espectral puede hacerse de manera externa, y puede almacenarse por el dispositivo de almacenamiento posteriormente. Por lo tanto, la complejidad computacional de un aparato realizado de acuerdo con el concepto descrito, puede reducirse de
40 manera significativa en comparación con los dispositivos conocidos, puesto que un cálculo explícito del factor de ponderación espectral no es necesario.

[0049] Los factores de ponderación espectral pueden calcularse para ser almacenados por el dispositivo de almacenamiento, por ejemplo, de la siguiente manera.

50 **[0050]** El procesamiento del audio puede realizarse en el dominio digital. En consecuencia, la señal de audio de entrada puede denotarse por la secuencia de tiempo discreto $x[n]$ que se ha muestreado de la fuente de audio a alguna frecuencia de muestreo f_s . Puede suponerse que la secuencia $x[n]$ se ha escalado de manera apropiada, de manera que la potencia rms de $x[n]$ en decibelios proporcionada por
55

$$60 \quad RMS_{dB} = 10 \log_{10} \left(\frac{1}{L} \sum_{n=0}^L x^2[n] \right)$$

es igual al nivel de presión del sonido en dB, a la cual el audio se está oyendo por un oyente humano. Además, la señal de audio puede suponerse como monofásica para simplicidad de la exposición.

65

[0051] La señal de audio de entrada se aplica a un banco de filtros de análisis o función de banco de filtros ("Banco de Filtros d Análisis"). Cada filtro en el Banco de Filtros de Análisis se diseña para simular la respuesta a la frecuencia en una ubicación particular a lo largo de la membrana basilar en el oído interno. El Filtro de Bancos puede incluir un conjunto de filtros lineales, cuyo ancho de banda y separación son constantes en la escala de frecuencia del Ancho de Banda Rectangular Equivalente (ERB), definida por Moore, Glasberg y Baer ("B. C. J. Moore, B. Glasberg, T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness", "supra").

[0052] Aunque la escala de la frecuencia ERB corresponde más estrechamente a la percepción humana y muestra el desempeño mejorado para producir mediciones del volumen objetivo que corresponden con los resultados del volumen subjetivo, la escala de frecuencia Bark puede emplearse con un desempeño reducido.

[0053] Para una frecuencia central f en Hertz, el ancho de una banda ERB en Hertz puede ser aproximadamente:

$$ERB(f) = 24.7(4.37f/1000 + 1) \tag{1}$$

[0054] De esta relación, se define una escala de la frecuencia deformada, de manera que en cualquier punto a lo largo de la escala deformada, la ERB correspondiente en unidades de la escala deformada, es igual a uno. La función para convertir de la frecuencia lineal en Hertz a esta escala de frecuencia ERB, se obtiene integrando el recíproco de la Ecuación 1:

$$\begin{aligned} Hz \text{ a } ERB(f) &= \int \frac{1}{24.7\left(\frac{4.37f}{1000} + 1\right)} df \\ &= 21.4 \log_{10}\left(\frac{4.37f}{1000} + 1\right) \end{aligned} \tag{2a}$$

[0055] También es útil expresar la transformación de la escala ERB nuevamente a la escala de frecuencia lineal, solucionando la Ecuación 2a para f :

$$ERB \text{ a } Hz(e) = f = \frac{1000}{4.37} 10^{(e/21.4-1)}, \tag{2b}$$

en donde e está en unidades de la escala ERB.

[0056] El Banco de Filtros de Análisis puede incluir B filtros auditivos, referidos como subbandas, a las frecuencias centrales $f_c[1] \dots f_c[B]$, separadas de manera uniforme a lo largo de la escala ERB. De manera más específica,

$$f_c[1] = f_{\min} \tag{3a}$$

para $b = 2 \dots B$

$$f_c[b] = f_c[b-1] + ERB \text{ a } Hz (Hz \text{ a } ERB(f_c[b-1]) + \Delta) \tag{3b}$$

$$f_c[B] < f_{\max}, \tag{3c}$$

en donde Δ es la separación ERB deseada del Banco de Filtros de Análisis, y en donde f_{\min} y f_{\max} son las frecuencias centrales mínima y máxima deseadas, respectivamente. Uno puede elegir $\Delta = 1$, y tomar en cuenta el intervalo de

frecuencia sobre el cual el oído humano es sensible, uno puede ajustar a $f_{\min} = 50$ Hz y $f_{\max} = 20.000$ Hz. Con tales parámetros, por ejemplo, la aplicación de las Ecuaciones 3a-c, proporciona $B = 40$ filtros auditivos.

5 **[0057]** La magnitud de la respuesta a la frecuencia para cada filtro auditivo puede representarse por una función exponencial redondeada, como se sugiere por Moore y Glasberg. De manera específica, la respuesta en magnitud de un filtro con la frecuencia central $f[b]$, puede calcularse como:

$$10 \quad H_b(f) = (1 + pg)e^{-pg} \quad (4a)$$

en donde

$$15 \quad g = \left| \frac{f - f_c[b]}{f_c[b]} \right|, \quad (4b)$$

$$20 \quad p = \frac{4f_c[b]}{ERB(f_c[b])} \quad (4c)$$

25 **[0058]** Las operaciones de filtrado del Banco de Filtros de Análisis puede aproximarse de manera adecuada utilizando una Transformada de Fourier Discreta de longitud finita, referida comúnmente como la Transformada de Fourier Discreta de Corto Plazo (STDFT), debido a que se cree que una implementación que se ejecuta en los filtros a la velocidad de muestreo de la señal de audio, referida como la implementación de velocidad completa, proporciona más resolución temporal de la necesaria para las mediciones exactas del volumen.

30 **[0059]** La STDFT de la señal de audio de entrada $x[n]$ puede definirse como:

$$35 \quad X[k, t] = \sum_{n=0}^{N-1} w[n]x[n + tT]e^{-j\frac{2\pi k}{N}n}, \quad (5a)$$

40 en donde k es el índice de la frecuencia, t es el índice del bloque de tiempo, N es el tamaño de la DFT, T es el tamaño del salto y $w[n]$ es una ventana de longitud N normalizada, de manera que

$$45 \quad \sum_{n=0}^{N-1} w^2[n] = 1 \quad (5b)$$

50 **[0060]** Nótese que la variable t en la Ecuación 5a es un índice discreto que representa el bloque de tiempo del STDFT en oposición a medir el tiempo en segundos. Cada incremento en t representa un salto de T muestras a lo largo de la señal $x[n]$. Las referencias posteriores al índice t suponen esta definición. Aunque pueden utilizarse diferentes ajustes del parámetro y formas de la ventana, dependiendo de los detalles de la implementación, para $f_s = 44100$ Hz, la elección de $N = 2048$, $T = 1024$, y hacer que $w[n]$ sea una ventana Hann, proporciona un equilibrio adecuado de la resolución de tiempo y la frecuencia. La STDFT descrita anteriormente puede implementarse de manera más eficiente utilizando la Transformada de Fourier Rápida (FFT).

60 **[0061]** En lugar de la STDFT, la Transformada del Coseno Discreto Modificada (MDCT) puede utilizarse para implementar el banco de filtros de análisis. La MDCT es una transformada utilizada comúnmente en los codificadores de audio perceptuales. La MDCT de la señal de audio de entrada $x[n]$ puede proporcionarse por:

$$65 \quad X[k, t] = \sum_{n=0}^{N-1} w[n]x[n + rT] \cos(2\pi/N)(k + 1/2)(n + n_0),$$

5 en donde

$$n_0 = \frac{(N/2) + 1}{2}$$

(6)

15 Generalmente, el tamaño del salto T se elige para ser exactamente la mitad de la longitud N de la transformada, de manera que es posible una reconstrucción perfecta de la señal x[n].

20 **[0062]** Las salidas del Banco de Filtros de Análisis se aplican a un filtro de transmisión o a una función del filtro de transmisión ("Filtro de Transmisión"), que filtra cada banda del banco de filtros de acuerdo con la transmisión del audio a través del oído externo y medio.

25 **[0063]** Con el fin de calcular el volumen de la señal de audio de entrada, se necesita una medición de la energía de corto plazo de las señales de audio en cada filtro del Banco de Filtros de Análisis después de la aplicación del Filtro de Transmisión. Esta medición variable del tiempo y la frecuencia se refiere como la excitación. La salida de energía de corto plazo de cada filtro en el Banco de Filtros de Análisis, a, puede aproximarse en una función de Excitación E[b,t] a través de la multiplicación de las respuestas del filtro en el dominio de la frecuencia con el espectro de la potencia de la señal de entrada:

$$E[b, t] = \frac{1}{N} \sum_{k=0}^{N-1} |H_b[k]|^2 |P[k]|^2 |X[k, t]|^2,$$

35 en donde b es en número de la subbanda, t es el número de bloque, y H_b[k] y P[k] son las respuestas a la frecuencia del filtro auditivo y el filtro de transmisión, respectivamente, muestreados a una frecuencia que corresponde a STDFT o MDCT b en el índice k. Deberá notarse que las formas para la respuesta de la magnitud de los filtros auditivos, diferentes a aquéllas especificadas en las Ecuaciones 4a-c, pueden utilizarse en la Ecuación 7 para lograr resultados similares.

40 **[0064]** En resumen, la salida de la Función de Excitación es una representación del dominio de la frecuencia de la energía E en las bandas ERB b respectivas por periodo de tiempo t.

45 **[0065]** Para ciertas aplicaciones, puede ser deseable homogeneizar la excitación E[b,t] antes de su transformación al volumen específico. Por ejemplo, la homogeneización puede realizarse de manera recursiva en una función de Homogeneización de acuerdo con la ecuación:

$$\bar{E}[b,t] = \lambda_b \bar{E}[b,t] + (1 - \lambda_b) E[b,t],$$

(8)

50 en donde las constantes de tiempo λ_b en cada banda b, se seleccionan de acuerdo con la aplicación deseada. En la mayoría de los casos, las constantes de tiempo pueden elegirse de manera ventajosa para ser proporcionales al tiempo de integración de la percepción del volumen humano dentro de la banda b. Watson y Gengel realizaron experimentos que demuestran que este tiempo de integración está dentro del intervalo de 150-175 ms a bajas frecuencias (125-200 Hz) y 40-60 ms a altas frecuencias ("Charles S. Watson y Roy W. Gengel, "Signal Duration and Signal Frequency in Relation to Auditory Sensitivity" Journal of the Acoustical Society of America, Vol. 46, No. 4 (Parte 2), 1969, pp. 989-997").

60 **[0066]** En una función de conversión ("Volumen Específico"), cada banda de la frecuencia de la excitación puede convertirse en un valor del componente del volumen específico, que se mide en Sone por ERB.

65 **[0067]** Inicialmente, al calcular el volumen específico, el nivel de excitación en cada banda de $\bar{E}[b,t]$ puede transformarse a un nivel de excitación equivalente a 1 kHz como se especifica mediante, por ejemplo, los contornos del volumen igual normalizados por un filtro de transmisión:

$$\bar{E}_{1kHz}[b,t] = T_{1kHz}(\bar{E}[b,t], f_c[b]),$$

(9)

en donde $T_{1kHz}(E, f)$ es una función que genera el nivel a 1 kHz, que es igualmente fuerte al nivel E a la frecuencia f . La transformación a los niveles equivalentes a 1 kHz simplifica el siguiente cálculo del volumen específico.

[0068] A continuación, el volumen específico en cada banda puede calcularse como:

$$N[b,t] = \alpha[b,t]N_{NB}[b,t] + (1 - \alpha[b,t])N_{WB}[b,t],$$

(10)

en donde $N_{NB}[b,t]$ y $N_{WB}[b,t]$ son los valores del volumen específico basados en un modelo de la señal de banda estrecha y de banda ancha, respectivamente. El valor $\alpha[b,t]$ es un factor de interpolación que cae entre 0 y 1, que es calculado de la señal de audio.

[0069] Los valores del volumen específico de banda estrecha y banda ancha $N_{NB}[b,t]$ y $N_{WB}[b,t]$ pueden estimarse de la excitación transformada utilizando las funciones exponenciales:

$$N_{NB}[b,t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b,t]}{TQ_{1kHz}} \right)^{\beta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b,t] > TQ_{1kHz} \\ 0, & \text{en caso contrario} \end{cases}$$

$$N_{WB}[m,t] = \begin{cases} G_{WB} \left(\left(\frac{\bar{E}_{1kHz}[b,t]}{TQ_{1kHz}} \right)^{\beta_{WB}} - 1 \right), & \bar{E}_{1kHz}[b,t] > TQ_{1kHz} \\ 0, & \text{en caso contrario} \end{cases}$$

(11a, 11b)

en donde TQ_{1kHz} es el nivel de excitación en el umbral en silencio para un tono de 1 kHz. De los contornos del volumen igual TQ_{1kHz} es igual a 4.2 dB. Una nota que ambas de estas funciones del volumen específico son iguales a cero cuando la excitación es igual al umbral en silencio. Para excitaciones mayores que el umbral en silencio, ambas funciones crecen de manera monótonica con una ley de potencia de acuerdo con la ley de Stevens de sensación de la intensidad. El exponente para la función de la banda estrecha se elige para ser mayor que aquél de la función de la banda ancha, haciendo que la función de la banda estrecha se incremente más rápidamente que la función de la banda ancha. La selección específica de los exponentes β y las ganancias G para los casos de la banda estrecha y la banda ancha se eligen para corresponder con los datos experimentales en el crecimiento del volumen para los tonos y ruidos.

[0070] El volumen específico puede ser igual a algún valor pequeño en lugar de cero cuando la excitación está en el umbral del oído. El volumen específico debe disminuir entonces de manera monótonica a cero conforme la excitación disminuye a cero. La justificación es que el umbral del oído es un umbral probabilístico (el punto en el cual un tono se detecta 50% del tiempo), y que varios tonos, cada uno en el umbral, presentados juntos, pueden sumar un sonido que es más audible que cualquiera de los tonos individuales. Si el volumen específico se define como cero cuando la excitación está en, o por debajo del umbral, entonces no existe una solución única para el solucionador de la ganancia para las excitaciones en, o por debajo del umbral. Si, por otra parte, el volumen específico se define como que se incrementa de manera monótonica para todos los valores de excitación mayores que, o iguales a cero, entonces existe una solución única. El escalar el volumen más que la unidad, siempre

resultará en una ganancia mayor que la unidad y viceversa. Las funciones del volumen específico en las Ecuaciones 11a y 11b pueden alterarse para tener la propiedad deseada de acuerdo con:

$$N_{NB}[b, t] = \begin{cases} G_{NB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\beta_{NB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{NB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{NB} \right)^{\eta_{NB}} \right\}, & \text{en caso contrario} \end{cases}$$

$$N_{WB}[m, t] = \begin{cases} G_{WB} \left(\left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right)^{\beta_{WB}} - 1 \right), & \bar{E}_{1kHz}[b, t] > \lambda TQ_{1kHz} \\ \exp \left\{ K_{WB} \left(-\log \left(\frac{\bar{E}_{1kHz}[b, t]}{TQ_{1kHz}} \right) + C_{WB} \right)^{\eta_{WB}} \right\}, & \text{en caso contrario} \end{cases}$$

(11c, 11d)

en donde la constante λ , es mayor que uno, el exponente η es menor que uno, y las constantes K y C se eligen de manera que la función del volumen específico y su primera derivada sean continuas en el punto:

$$\bar{E}_{1kHz}[b, t] = \lambda TQ_{1kHz}$$

[0071] Del volumen específico, el volumen general o "total" $L[t]$, se proporciona por la suma del volumen específico a través de todas las bandas b:

$$L[t] = \sum_b N[b, t]$$

(12)

[0072] En una función de modificación del volumen específico ("Modificación del Volumen Específico"), el volumen específico objetivo, referido como $\hat{N}[b, t]$, puede calcularse del volumen específico de varias formas. Como se describe con mayor detalle a continuación, un volumen específico objetivo puede calcularse utilizando un factor de escala α , por ejemplo, en el caso de un control del volumen. Véase la Ecuación 16 a continuación, y su descripción asociada. En el caso de un control de ganancia automático (AGC) y un control del intervalo dinámico (DRC), un volumen específico objetivo puede calcularse utilizando una relación del volumen de salida deseado al volumen de entrada. Véanse las Ecuaciones 17 y 18 a continuación y sus descripciones asociadas. En el caso de una igualación dinámica, un volumen específico objetivo puede calcularse utilizando una relación expuesta en la Ecuación 23 y su descripción asociada.

[0073] En este ejemplo, para cada banda b y cada intervalo de tiempo t, una función de solución de la ganancia toma como su entrada la excitación $\bar{E}[b, t]$ homogeneizada y el volumen específico objetivo $\hat{N}[b, t]$, y genera los factores de ponderación espectral, también llamados ganancias $G[b, t]$, utilizados posteriormente para modificar el

audio. Sea la función $\psi\{\cdot\}$ que representa la transformación no lineal de la excitación al volumen específico, de manera que

$$N[b,t]=\Psi\{\bar{E}[b,t]\}, \tag{13}$$

el Solucionador de la Ganancia encuentra $G[b,t]$, de manera que

$$\hat{N}[b,t]=\Psi\{G^2[b,t]\bar{E}[b,t]\}. \tag{14a}$$

[0074] La función de solución de la ganancia determina las ganancias que varían con la frecuencia y con el tiempo (factores de ponderación espectral), que cuando se aplican a la excitación original, resultan en un volumen específico que, idealmente, es igual al volumen específico objetivo deseado. En la práctica, la función de solución de la ganancia determina las ganancias que varían con la frecuencia y el tiempo, que cuando se aplican a la versión del dominio de la frecuencia de la señal de audio, resulta en la modificación de la señal de audio, con el fin de reducir la diferencia entre su volumen específico y el volumen específico objetivo. Idealmente, la modificación es tal, que la señal de audio modificada tiene un volumen específico que es una aproximación de una dosis del volumen específico objetivo. La solución a la Ecuación 14a puede implementarse en una variedad de formas. Por ejemplo, si existe una expresión matemática de forma cerrada para el inverso del volumen específico, representado por $\psi^{-1}\{\cdot\}$, entonces las ganancias pueden calcularse directamente re-arreglando la ecuación 14a:

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\hat{N}[b, t])}{\bar{E}[b, t]}} \tag{14b}$$

Alternativamente, si no existe una solución de forma cerrada para $\psi^{-1}\{\cdot\}$, puede emplearse un procedimiento iterativo en el cual, para cada iteración, la ecuación 14a se evalúa utilizando un estimado actual de las ganancias. El volumen específico resultante se compara con el objetivo deseado y las ganancias se actualizan basándose en el error. Si las ganancias se actualizan de manera apropiada, convergerán a la solución deseada. Como se mencionó anteriormente, el volumen específico objetivo puede representarse por un escalamiento del volumen específico:

$$\hat{N}[b,t]=\Xi[b,t]N[b,t] \tag{14c}$$

[0075] Sustituyendo la ecuación 13 en 14c y a continuación 14c en 14b, se proporciona una expresión alterna para las ganancias:

$$G[b, t] = \sqrt{\frac{\Psi^{-1}(\Xi[b, t]\Psi(\bar{E}[b, t]))}{\bar{E}[b, t]}} \tag{14d}$$

[0076] Los factores de ponderación espectral calculados o ganancias se almacenan en la tabla de búsqueda del dispositivo de almacenamiento.

[0077] En algunas realizaciones de acuerdo con la invención, el determinador de la excitación no determina un valor de un parámetro de la excitación para todas las subbandas de la pluralidad de subbandas. En este caso, es suficiente que la tabla de búsqueda contenga sólo los factores de ponderación espectral asociados con las subbandas, para las cuales se determina un valor de un parámetro de la excitación. De esta manera, el espacio de almacenamiento del dispositivo de almacenamiento necesario para almacenar la tabla de búsqueda puede reducirse de manera significativa.

[0078] Puesto que la flexión de los contornos del volumen igual, que deben ser compensados, es más fuerte para las frecuencias más bajas (véanse las Figura 2 y 3), puede ser suficiente compensar una variación del volumen sólo para las subbandas de baja frecuencia. Por lo tanto, puede ser útil calcular los parámetros de la excitación y almacenar los factores de ponderación espectral para las subbandas de baja frecuencia. En contraste, para las subbandas de alta frecuencia, ningún valor de un parámetro de la excitación puede determinarse, y ningún factor de ponderación espectral asociado con las subbandas de alta frecuencia puede almacenarse. En otras palabras, una subbanda, para la cual se determina un valor de un parámetro de la excitación, puede comprender frecuencias más bajas que una subbanda para la cual se determina un valor de un parámetro de la excitación.

[0079] Además, puede no ser necesario modificar las subbandas de alta frecuencia. En otras palabras, un contenido de una subbanda puede no modificarse por el modificador de la señal, si el determinador de la excitación no determina un valor de un parámetro de la excitación para esta subbanda. Este puede ser sólo el caso, si no se considera ningún otro parámetro, tal como, por ejemplo, un parámetro de modificación externa o un parámetro del ruido de fondo.

[0080] Alternativamente, un factor de ponderación espectral proporcionado por el dispositivo de almacenamiento, puede utilizarse por el modificador de la señal para más de una subbanda. En otras palabras, el modificador de la señal puede modificar un contenido de una subbanda, para la cual no se determina un valor de un parámetro de la excitación, basándose en un factor de ponderación espectral proporcionado para una subbanda, para la cual se determina un valor de un parámetro de la excitación. Considerando el comportamiento de los contornos del volumen igual mostrados en las Figuras 2 y 3, puede ser suficiente modificar las bandas de alta frecuencia de acuerdo con el mismo factor de ponderación espectral. Este factor de ponderación espectral puede ser el factor de ponderación espectral proporcionado para la subbanda que comprende las frecuencias más altas de todas las subbandas, para las cuales se determina un valor de un parámetro de la excitación. Más generalmente, el modificador de la señal puede modificar un contenido de una subbanda, para la cual no se determina un valor de un parámetro de la excitación, basándose en el factor de ponderación espectral proporcionado para una subbanda que contiene frecuencias más altas que todas las otras subbandas, para las cuales se determina un valor de un parámetro de la excitación. Por ejemplo, puede ser suficiente que el determinador de la excitación determine el valor de un parámetro de la excitación sólo para 5 a 15 (o 2 a 20, 7 a 12 o sólo 5, 6, 7, 8, 9, 10, 11, 12) subbandas de la pluralidad de subbandas, o sólo para menos de un cuarto, un tercio, la mitad o dos tercios de las subbandas de la pluralidad de subbandas. Estas subbandas pueden comprender frecuencias más bajas que todas las otras subbandas de la pluralidad de subbandas. Además, el modificador de la señal puede modificar el contenido de estas subbandas de acuerdo con los factores de ponderación espectral proporcionados por el dispositivo de almacenamiento para estas subbandas.

[0081] Por ejemplo, la escala Bark comprende 25 bandas de frecuencia y puede ser suficiente para modificar las 7 bandas de frecuencia más baja, puesto que las bandas con la frecuencia más baja muestran la desviación más fuerte del comportamiento inactivo. De manera alterna, las bandas más bajas de la escala ERB pueden modificarse. Las subbandas restantes de la pluralidad de subbandas pueden permanecer no modificadas, pueden modificarse de acuerdo con un parámetro de modificación externa y/o un parámetro del ruido de fondo, o puede modificarse de acuerdo con el factor de ponderación espectral proporcionado para una subbanda, para la cual se determina un valor de un parámetro de la excitación, que contiene las frecuencias más altas que todas las otras subbandas, para las cuales se determina un valor de un parámetro de la excitación.

[0082] La Figura 5 muestra un diagrama de flujo de un método 500 para modificar una señal de audio de entrada de acuerdo con una realización de la invención. El método 500 comprende determinar en 510 un valor de un parámetro de la excitación de una subbanda de una pluralidad de subbandas de la señal de audio de entrada, basándose en un contenido de energía de la subbanda. Además, el método 500 comprende proporcionar en 520 un factor de ponderación espectral que corresponde al valor determinado del parámetro de la excitación, y que corresponde a la subbanda, para la cual se determina el valor del parámetro de la excitación. El factor de ponderación espectral se almacena en una tabla de búsqueda que contiene una pluralidad de factores de ponderación espectral. Un factor de ponderación espectral de la pluralidad de factores de ponderación espectral se asocia con un valor predefinido del parámetro de la excitación y una subbanda de la pluralidad de subbandas. Finalmente, el método 500 comprende modificar la subbanda, para la cual se determina el valor del parámetro de la excitación, basándose en el factor de ponderación espectral proporcionado para obtener y proporcionar una subbanda modificada.

[0083] En otras palabras, el método 500 comprende un cálculo en 510 de una señal de excitación, recuperando en 520 las ponderaciones espectrales (factores de ponderación espectral) de la tabla de búsqueda y modificar en 530 la

señal de audio de salida. Opcionalmente, el método 500 comprende una re-síntesis de la señal de audio de salida (combinando las subbandas para obtener una señal de audio modificada).

[0084] Esto puede, por ejemplo, ser un método para una modificación de la señal eficiente y genérica.

[0085] Además opcionalmente, un parámetro de modificación externa también puede tomarse en cuenta (indicado por la línea punteada) como se describió anteriormente.

[0086] Una consideración adicional del nivel de la subbanda del ruido de fondo (un parámetro del ruido de fondo), se menciona por el método 600 mostrado en la Figura 6.

[0087] Algunas realizaciones de acuerdo con la invención, se relacionan con una realización eficiente del procesamiento perceptual de las señales de audio. El concepto descrito se relaciona con una arquitectura flexible y altamente eficiente para la modificación y procesamiento de la señal de audio selectiva con la frecuencia, que puede incorporar fácilmente las características de los efectos psicoacústicos en su procesamiento, sin sufrir de la carga computacional del modelado auditivo explícito. Como un ejemplo, se considera que la realización de un procesador con múltiples bandas para el control del volumen perceptual se basa en la arquitectura mostrada.

[0088] Esta puede ser una realización eficiente del control del volumen psicoacústico.

[0089] El procesamiento descrito anteriormente es comparable con la filtración de la señal de entrada con un filtro característico, que es controlado por el nivel de entrada dentro de cada banda de frecuencia auditiva. Puede implementarse de manera más eficiente.

[0090] Básicamente, el método propuesto evita el cálculo del volumen específico y el cálculo hacia atrás correspondiente, y por lo tanto, evita los pasos de procesamiento computacional intensos a costo de requisitos de memoria incrementados ligeramente.

[0091] La implementación eficiente puede implementarse utilizando una tabla de búsqueda simple (LUT), posiblemente con interpolación.

[0092] La LUT se calcula midiendo los valores de entrada y los valores de salida del proceso implementado como se describió anteriormente. La LUT tiene, por ejemplo, 3 dimensiones. Produce una subbanda modificada o una señal de audio modificada dada la excitación de entrada, el parámetro de modificación y el índice de la banda de la frecuencia.

[0093] Por ejemplo, puede implementarse de manera eficiente reconociendo que su funcionalidad es dependiente del índice de la banda de la frecuencia sólo para las bandas de la frecuencia más baja, por ejemplo, cuando se utiliza un banco de filtros auditivos con una resolución que corresponde a la escala Bark, el banco de filtros puede tener 25 filtros de paso de banda. El almacenamiento de la función de transferencia en la LUT para las 7 bandas más bajas únicamente, puede ser suficiente, puesto que para los índices de las bandas superiores, la misma relación de entrada-salida se mantiene para el índice de la banda 7.

[0094] Este procesamiento eficiente proporciona un control del volumen que es correcto en un sentido psicoacústico. Otras aplicaciones, a saber el control del intervalo dinámico y/o la igualación dinámica, se derivan del procesamiento eficiente descrito anteriormente, indexando de manera apropiada la LUT.

[0095] Finalmente, la compensación del ruido de fondo (es decir, la compensación del efecto del enmascaramiento parcial de una señal de audio en la presencia del ruido de fondo), puede lograrse agregando una cuarta dimensión a la LUT, que representa el nivel del ruido de fondo. El diagrama de bloques del procesamiento propuesto para la compensación del ruido se ilustra en la Figura 6.

[0096] Aunque el procesamiento descrito hasta ahora estaba dirigido a la emulación de un algoritmo de escalamiento del volumen psicoacústico, la arquitectura descrita en la Figura 1 o la Figura 4 puede producir un espectro mucho más rico de modificaciones del sonido que estarían disponibles con un algoritmo de escalamiento del volumen psicoacústico desde su LUT. Puede hacerse dependiente de aún más factores (por ejemplo, un ajuste de la preferencia del usuario, otros factores que varían con el tiempo, etc.). Puede "afinarse" libremente de acuerdo con la preferencia subjetiva del oyente más allá de las características que se proporcionan por una función dada como una expresión de forma cerrada.

[0097] En resumen, la invención se relaciona con una arquitectura flexible y altamente eficiente para la modificación y procesamiento de la señal de audio, selectiva de la frecuencia, que puede incorporar fácilmente las características de los efectos psicoacústicos en su procesamiento, sin sufrir de la carga computacional del modelado auditivo explícito.

- 5 **[0098]** A un nivel abstracto, el procesamiento eficiente propuesto comprende los siguientes pasos. Basándose en la señal de entrada, uno o más valores de las características (incluyendo el valor del parámetro de la excitación) pueden calcularse para varias bandas de frecuencia (por ejemplo, bandas críticas). Basándose en estos valores de las características (y posiblemente, otra información), se realiza una tabla de búsqueda para cada una de estas
- 10 **[0099]** El procesamiento de las señales de audio en las bandas de frecuencia implica usualmente utilizar bancos de filtros, es decir, la señal de entrada es dividida en varias bandas de frecuencia (subbandas) por un banco de filtros de análisis, y la señal de salida final se obtiene alimentando las señales de la subbanda modificadas en el banco de filtros de síntesis. Los bancos de filtros de análisis y de síntesis se combinan para reconstruir la señal del tiempo de entrada, ya sea perfectamente o casi perfectamente.
- 15 **[0100]** Un número típico de bandas de frecuencia es de entre 4 y 40. La tabla de búsqueda basada en los valores de las características involucra usualmente la cuantificación de los valores de las características en un conjunto limitado de valores, que pueden utilizarse como un índice de búsqueda en la tabla. Además, el tamaño de la tabla de búsqueda puede reducirse eligiendo un tamaño del paso de cuantificación muy aproximado y posteriormente interpolando entre los (dos o más) valores adyacentes del parámetro de salida de la tabla. Con el fin de considerar
- 20 varias características de entrada para el cálculo de los valores del parámetro de salida, puede utilizarse una tabla de búsqueda con varias dimensiones, por ejemplo, la LUT del factor de modificación que contiene el idx (índice) de excitación, idx de la tonalidad, idx de la frecuencia. En un caso muy simple (y eficiente), los valores del parámetro de salida representan directamente los factores de multiplicación a ser aplicados a la subbanda de entrada, con el fin de
- 25 determinar las señales de la subbanda de salida. Esto se muestra, por ejemplo, en la Figura 4.
- 30 **[0101]** Aunque algunos aspectos del concepto descrito se han explicado en el contexto de un aparato, está claro que estos aspectos también representan una descripción del método correspondiente, en donde un bloque o dispositivo corresponde a un paso del método o a una característica de un paso del método. De manera análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o punto o característica correspondiente, de un aparato correspondiente.
- 35 **[0102]** Dependiendo de ciertos requisitos de la implementación, las realizaciones de la invención pueden implementarse en los elementos físicos o en el programa. La implementación puede realizarse utilizando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en la misma, que cooperan (o son capaces de cooperar) con un sistema de computadora programable, de manera que se realiza el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por
- 40 computadora.
- 45 **[0103]** Algunas realizaciones de acuerdo con la invención, comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema de computadora programable, de manera que se realiza uno de los métodos descritos en la presente.
- 50 **[0104]** Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de un programa para computadora con un código del programa, el código del programa es operativo para realizar uno de los métodos cuando el producto de un programa para computadora se ejecuta en una computadora. El código del programa puede, por ejemplo, almacenarse en un portador legible por una máquina.
- 55 **[0105]** Otras realizaciones comprende el programa para computadora para realizar uno de los métodos descritos en la presente, almacenado en un portador legible por una máquina.
- 60 **[0106]** En otras palabras, una realización del método inventivo es, por lo tanto, un programa para computadora que tiene un código del programa para realizar uno de los métodos descritos en la presente, cuando el programa para computadora se ejecuta en una computadora.
- 65 **[0107]** Una realización adicional de los métodos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, registrado en el mismo, el programa para computadora para realizar uno de los métodos descritos en la presente.
- [0108]** Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa para computadora para realizar uno de los métodos descritos en la presente. El flujo de datos o la secuencia de señales, puede, por ejemplo, configurarse para transferirse vía una conexión de comunicación de datos, por ejemplo, vía la Internet.

[0109] Una realización adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado para, o adaptado para realizar uno de los métodos descritos en la presente.

5 **[0110]** Una realización adicional comprende una computadora que tiene instalado en la misma, un programa para computadora para realizar uno de los métodos descritos en la presente.

10 **[0111]** En algunas realizaciones, puede utilizarse un dispositivo lógico programable (por ejemplo, un arreglo de compuerta programable en el campo), para realizar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, un arreglo de compuerta programable en el campo puede cooperar con un microprocesador para realizar uno de los métodos descritos en la presente. Generalmente, los métodos son realizados de manera preferida por un aparato de un elemento físico.

15 **[0112]** Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entenderá que las modificaciones y variaciones de los arreglos y los detalles descritos en la presente, serán evidentes para otros con experiencia en la técnica. Por lo tanto, la intención es estar limitado únicamente por el alcance de las reivindicaciones de la patente que siguen a continuación y no por los detalles específicos presentados por medio de la descripción y explicación de las realizaciones presentes.

20

REIVINDICACIONES

1. Un aparato (100) para modificar una señal de audio de entrada, que comprende:

un determinador de la excitación (110) configurado para determinar un valor (112) de un parámetro de la excitación de una subbanda (102) de una pluralidad de subbandas de la señal de audio de entrada, basándose en un contenido de energía de la subbanda (102), en donde el valor (112) del parámetro de la excitación indica una potencia de la señal de audio en la subbanda o una energía de corto plazo de la señal de audio en la subbanda o un valor cuantificado de la energía de corto plazo de la señal de audio en la subbanda;

un dispositivo de almacenamiento (120) que almacena una tabla de búsqueda que contiene una pluralidad de factores de ponderación espectral, en donde un factor de ponderación espectral de la pluralidad de factores de ponderación espectral está asociado con un valor predefinido del parámetro de la excitación y con una subbanda de la pluralidad de subbandas, en donde el dispositivo de almacenamiento está configurado para proporcionar un factor de ponderación espectral (124) que corresponde al valor determinado (112) del parámetro de la excitación y que corresponde a la subbanda (102), para la cual se determina el valor (112) del parámetro de la excitación; y

un modificador de la señal (130) configurado para modificar un contenido de la subbanda (102) de la señal de audio de entrada, para la cual se determina el valor (112) del parámetro de la excitación, basándose en el factor de ponderación espectral (124) proporcionado, para proporcionar una subbanda modificada (132), mediante el escalado multiplicativo de la subbanda de la señal de audio por el factor de ponderación espectral proporcionado por la tabla de búsqueda.

2. El aparato según la reivindicación 1, en donde el determinador de la excitación (110) está configurado para determinar un valor (112) de un parámetro de la excitación para más de una subbanda (102) de la pluralidad de subbandas, en donde el dispositivo de almacenamiento (120) está configurado para proporcionar un factor de ponderación espectral (124) para cada subbanda (102), para la cual se determina un valor (112) de un parámetro de la excitación, y en donde el modificador de la señal (130) está configurado para modificar un contenido de cada subbanda (102), para la cual se determina un valor (112) de un parámetro de la excitación, basándose en el factor de ponderación espectral (124) respectivo, correspondiente proporcionado.

3. El aparato según la reivindicación 1 ó 2, caracterizado porque comprende además:

un banco de filtros de análisis (410) configurado para separar la señal de audio de entrada en la pluralidad de subbandas; y

un banco de filtros de síntesis (420) configurado para combinar la pluralidad de subbandas que contienen al menos una subbanda modificada (132) para proporcionar una señal de audio modificada.

4. El aparato según cualquiera de las reivindicaciones 1 a 3, en donde cada factor de ponderación espectral contenido en la tabla de búsqueda está asociado con un valor predefinido del parámetro de la excitación y con una subbanda de la pluralidad de subbandas.

5. El aparato según cualquiera de las reivindicaciones 1 a 4, en donde las subbandas de la pluralidad de subbandas de la señal de audio de entrada están divididas de acuerdo a la escala ERB, la escala Bark u otra separación frecuencial, que imita la resolución frecuencial del oído humano.

6. El aparato según cualquiera de las reivindicaciones 1 a 5, en donde el determinador de la excitación (110) está configurado para determinar un valor (112) de un parámetro de la excitación, no para todas las subbandas de la pluralidad de subbandas, y en donde la tabla de búsqueda contiene sólo factores de ponderación espectral asociados con las subbandas, para las cuales se determina un valor de un parámetro de la excitación.

7. El aparato según la reivindicación 6, en donde una subbanda (102), para la cual se determina un valor (112) de un parámetro de la excitación, comprende frecuencias más bajas que una subbanda, para la cual no se determina un valor de un parámetro de la excitación.

8. El aparato según la reivindicación 6 ó 7, en donde un contenido de una subbanda de la señal de audio de entrada no se modifica por el modificador de la señal (130), si el determinador de la excitación (110) no determina un valor (112) de un parámetro de la excitación para esta subbanda.

9. El aparato de según cualquiera de las reivindicaciones 1 a 8, en donde el determinador de la excitación (110) está configurado para determinar un valor (112) de un parámetro de la excitación sólo para menos que un tercio de las subbandas de la pluralidad de subbandas, y en donde el modificador de la señal (130) está configurado para modificar un contenido de las subbandas, para las cuales se determina un valor de un parámetro de la excitación, basándose en un factor de ponderación espectral respectivo, correspondiente proporcionado, y en donde estas

subbandas comprende frecuencias más bajas que todas las otras subbandas de la pluralidad de subbandas, para las cuales se determina un valor de un parámetro de la excitación.

5 **10.** El aparato según cualquiera de las reivindicaciones 1 a 9, en donde el modificador de la señal (130) está configurado para modificar un contenido de una subbanda, para la cual no se determina un valor de un parámetro de la excitación, basándose en un factor de ponderación espectral (124) proporcionado para una subbanda (102), para la cual se determina un valor (112) de un parámetro de la excitación.

10 **11.** El aparato según la reivindicación 10, en donde el modificador de la señal (130) modifica un contenido de la subbanda, para la cual no se determina un valor de un parámetro de la excitación, basándose en un factor de ponderación espectral (124) proporcionado para una subbanda (102), para la cual se determina un valor (112) de un parámetro de la excitación, que contiene frecuencias más altas que todas las otras subbandas (102), para las cuales se determina un valor (112) de un parámetro de la excitación.

15 **12.** El aparato según cualquiera de las reivindicaciones 1 a 11, en donde un factor de ponderación espectral contenido en la tabla de búsqueda está asociado además con un valor predefinido de un parámetro de modificación externa, en donde el dispositivo de almacenamiento (120) está configurado para proporcionar un factor de ponderación espectral (124) que corresponde al valor (112) determinado del parámetro de la excitación de una subbanda (102), que corresponde a la subbanda (102), para la cual se determina el valor (112) de un parámetro de la excitación, y que corresponde a un valor del parámetro de modificación externa.

20 **13.** El aparato según la reivindicación 12, en donde la tabla de búsqueda comprende exactamente tres dimensiones asociadas con los valores predefinidos del parámetro de la excitación, con las subbandas de la pluralidad de subbandas y con los valores predefinidos del parámetro de modificación externa.

25 **14.** El aparato según cualquiera de las reivindicaciones 12 a 13, en donde el modificador de la señal (130) está configurado para modificar un contenido de una subbanda, para la cual no se determina un valor de un parámetro de la excitación, basándose en un valor del parámetro de modificación externa.

30 **15.** El aparato según cualquiera de las reivindicaciones 1 a 14, en donde un factor de ponderación espectral contenido en la tabla de búsqueda está asociado además con un valor predefinido de un parámetro del ruido de fondo, en donde el dispositivo de almacenamiento (120) está configurado para proporcionar un factor de ponderación espectral (124) que corresponde al valor (112) determinado del parámetro de la excitación de la subbanda (102), que corresponde a la subbanda (102), para la cual se determina el valor (112) de un parámetro de la excitación, y que corresponde a un valor del parámetro del ruido de fondo.

35 **16.** El aparato según la reivindicación 15, en donde la tabla de búsqueda comprende exactamente cuatro dimensiones asociadas con los valores predefinidos del parámetro de la excitación, con las subbandas de la pluralidad de subbandas, con los valores predefinidos del parámetro de modificación externa y con los valores predefinidos del parámetro del ruido de fondo.

40 **17.** El aparato según cualquiera de las reivindicaciones 1 a 16, en donde el dispositivo de almacenamiento (120) no comprende ninguna entrada para un parámetro del volumen específico o un parámetro del volumen específico objetivo.

45 **18.** El aparato según cualquiera de las reivindicaciones 1 a 17, en donde la tabla de búsqueda almacenada por el dispositivo de almacenamiento (120) es la única tabla de búsqueda del aparato para modificar la señal de audio de entrada.

50 **19.** El aparato según cualquiera de las reivindicaciones 1 a 18, en donde el determinador de la excitación (110) está configurado para medir un contenido de energía de la subbanda (102) y configurado para cuantificar el contenido de energía medido de la subbanda para obtener el valor del parámetro de la excitación, de manera que el valor del parámetro de la excitación es igual a un valor predefinido del parámetro de la excitación.

55 **20.** Un método (500, 600) para modificar una señal de audio de entrada, que comprende:

determinar (510) un valor de un parámetro de la excitación de una subbanda de una pluralidad de subbandas de la señal de audio de entrada, basándose en un contenido de energía de la subbanda, en donde el valor del parámetro de la excitación indica una potencia de la señal de audio en la subbanda o una energía de corto plazo de la señal de audio en la subbanda o un valor cuantificado de la energía de corto plazo de la señal de audio en la subbanda;

60 proporcionar (520) un factor de ponderación espectral que corresponde al valor determinado del parámetro de la excitación y que corresponde a la subbanda, para la cual se determina el valor de el parámetro de la excitación, en donde el factor de ponderación espectral está almacenado en una tabla de búsqueda que contiene una pluralidad de factores de ponderación espectral, en donde un factor de ponderación espectral de la pluralidad de factores de

ponderación espectral está asociado con un valor predefinido del parámetro de la excitación y con una subbanda de la pluralidad de subbandas;

5 modificar (530) la subbanda, para la cual se determina el valor del parámetro de la excitación, basándose en el factor de ponderación espectral proporcionado, para proporcionar una subbanda modificada mediante el escalado multiplicativo de la subbanda de la señal de audio por el factor de ponderación espectral proporcionado por la tabla de búsqueda.

10 **21.** Un programa para computadora con un código del programa para realizar el método según la reivindicación 20, cuando el programa para computadora se ejecuta en una computadora o un microcontrolador.

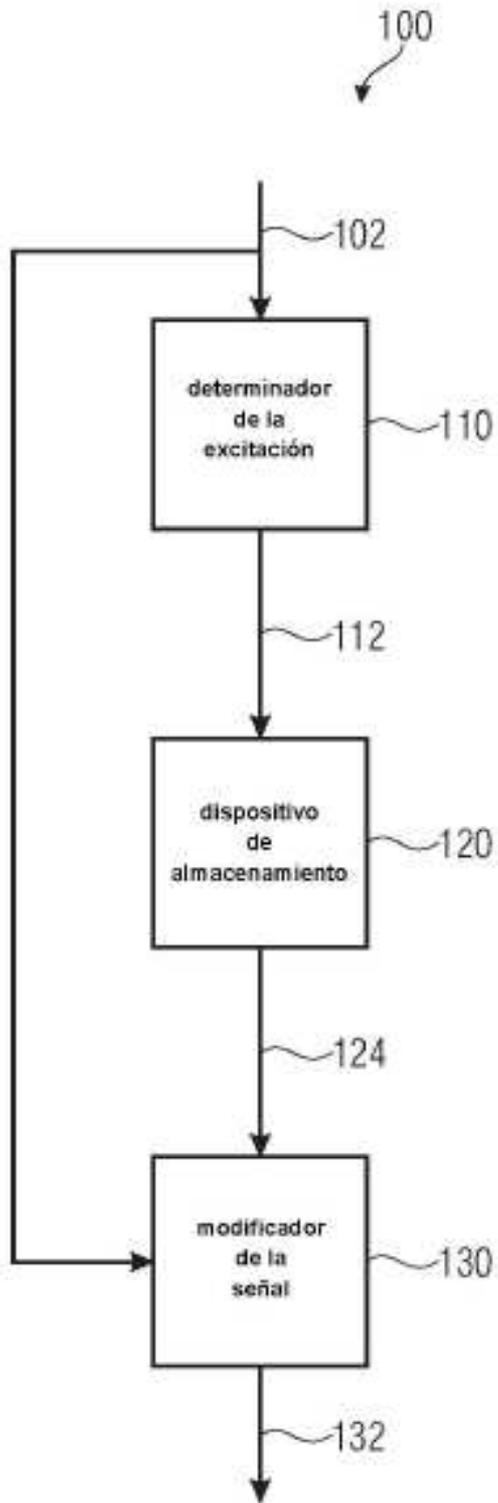


FIGURA 1

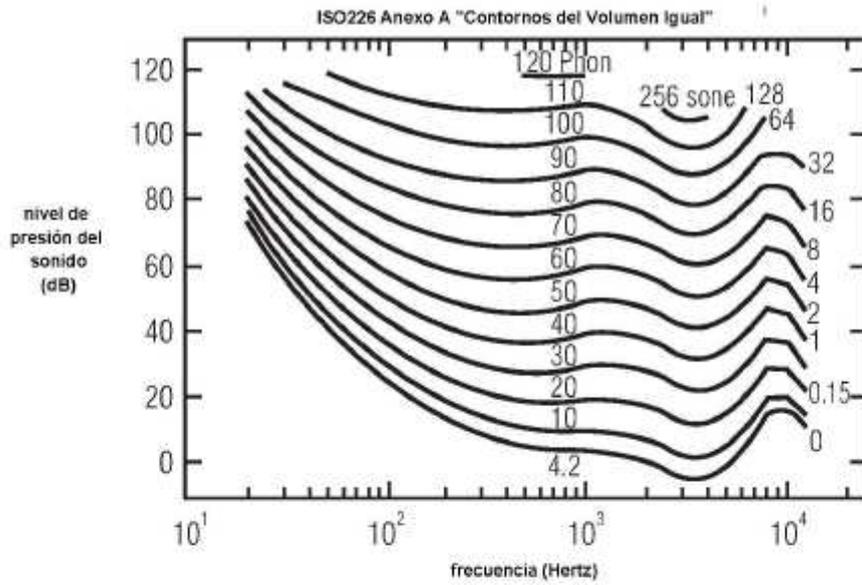


FIGURA 2

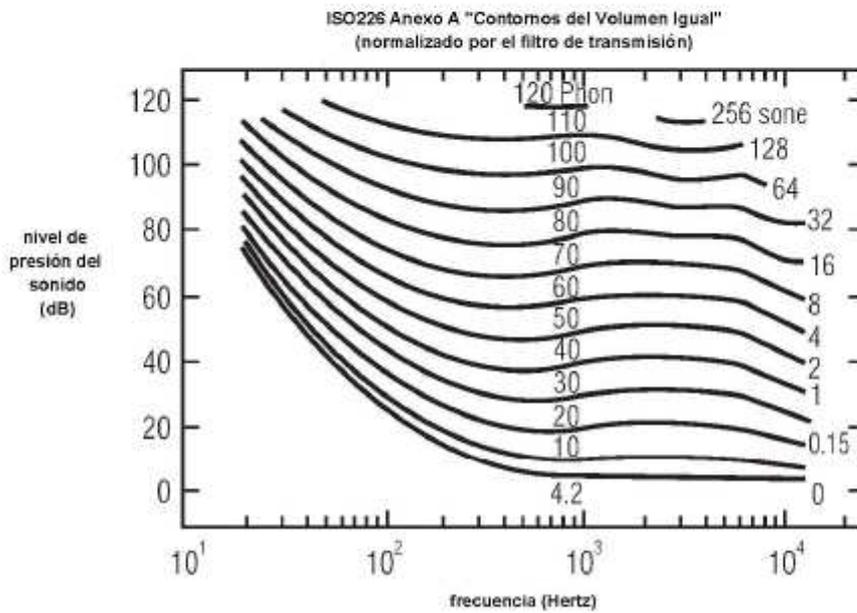


FIGURA 3

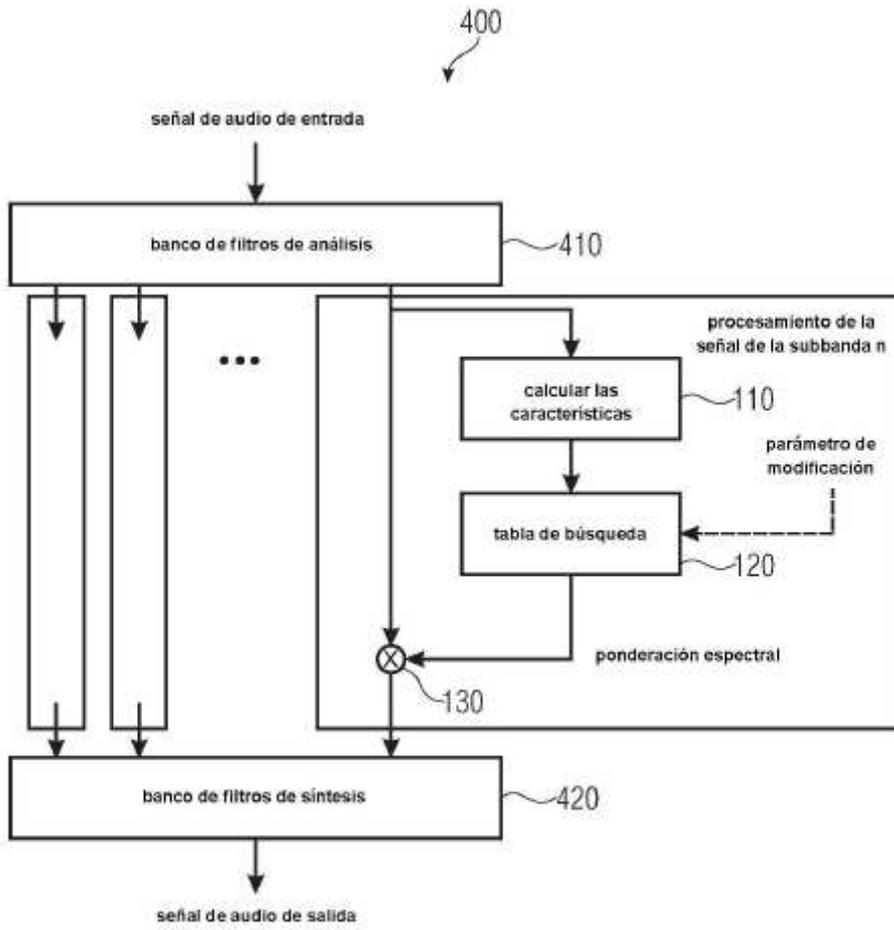


FIGURA 4

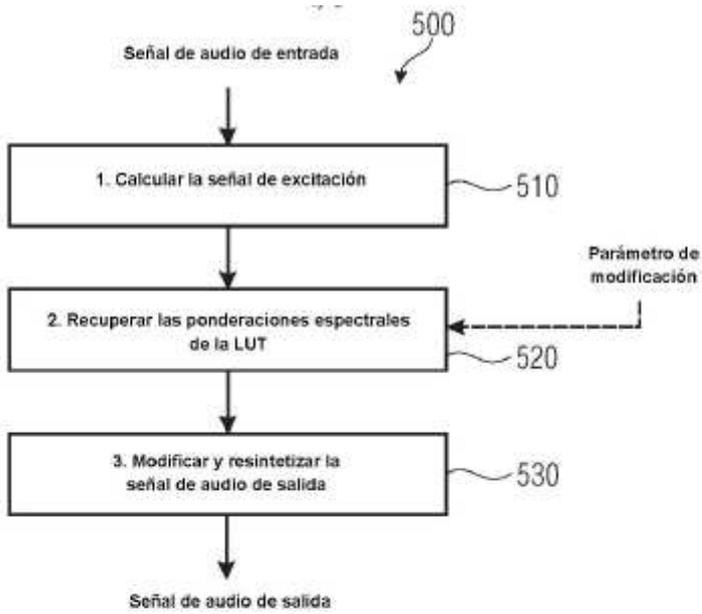


FIGURA 5

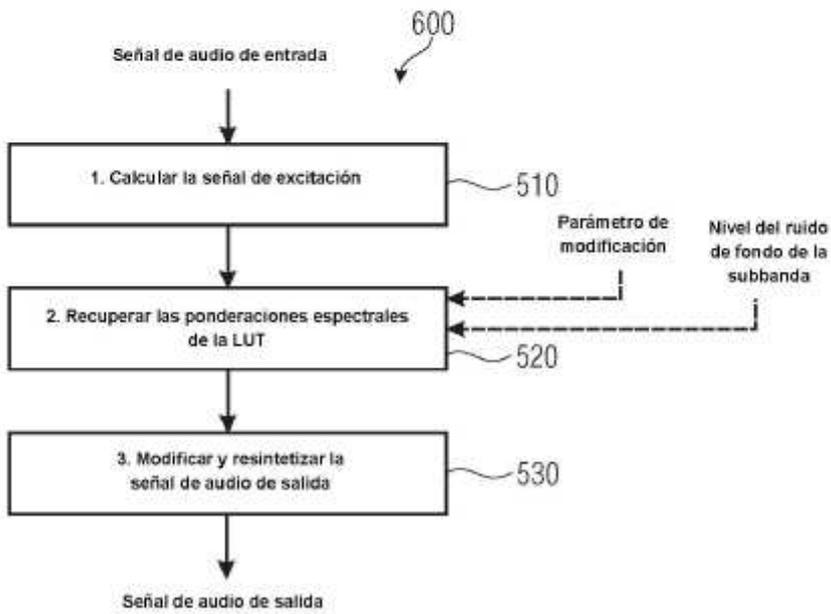


FIGURA 6

